

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:39:49 ; Search time 17 Seconds

(without alignments)
961.344 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893

Sequence: 1 VALNSPLSGMGRIGRADPQ.....ASCHNAVIVCIENSFMPTAS 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database : PIR 73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	684	2	A53019 collagen alpha 1(X
2	778	87.1	1774	2	B56101 collagen alpha 1(X
3	775	86.8	1315	2	A56101 collagen alpha 1(X
4	504	56.4	1388	2	A53317 collagen alpha 1(X
5	357	40.0	650	2	T22002 pectinesterase hom
6	79	8.8	477	2	T05202 pectinesterase hom
7	86	8.6	314	2	F98231 succinoglycan bios
8	77	8.6	314	2	AG3054 succinoglycan bios
9	77	8.5	204	2	B45022 CRK-I - human
10	76	8.5	636	2	JW0047 class I cyclinase
11	75.5	8.5	904	2	T04377 probable pullulan
12	75	8.4	351	2	B89781 conserved hypotet
13	74.5	8.3	850	2	SS6015 gastric mucin MUC5
14	73.5	8.2	1653	2	B91052 hypotetrical prote
15	73.5	8.2	1653	2	F88896 hypotetrical prote
16	73.5	8.2	1653	2	G65028 hypotetrical prote
17	73	8.2	435	2	JC5584 hyaluronoglycosam
18	73	8.2	462	2	T17480 endo-xyranase homo
19	73	8.2	513	1	A45333 exopolysphatase
20	73	8.2	513	2	D91049 exopolysphatase
21	73	8.2	513	2	H85893 exopolysphatase
22	72.5	8.1	884	2	C70729 hypotetrical prote
23	72	8.1	239	2	A46243 epidermal growth f
24	72	8.1	297	2	AB3128 carboxylate kinase
25	72	8.1	304	2	A45022 CRK-II - human
26	72	8.1	319	2	D98159 C-Crk - mouse
27	72	8.1	319	2	D98159 hypotetrical prote
28	72	8.1	800	2	AG0438 probable Rns acces
29	72	8.1	1006	2	T00050 hypotetrical prote

30	71.5	8.0	554	2	T06374 probable pectinest
31	71.5	8.0	554	2	T06468 pectinesterase (EC
32	71	8.0	579	2	A70954 hypotetrical prote
33	71	8.0	644	2	A72519 probable 2-oxoacid
34	71	8.0	732	2	AF0439 probable Rns acces
35	71	8.0	787	2	G81692 inner membrane pro
36	70.5	7.9	774	2	T04892 probable anthranil
37	70	7.8	375	2	C75368 conserved hypotet
38	70	7.8	375	2	F83906 unsaturated glucur
39	70	7.8	451	2	H75593 hypotetrical prote
40	70	7.8	672	2	T36083 hypotetrical prote
41	70	7.8	741	2	AC0094 conserved hypotet
42	70	7.8	777	2	AD0982 biotin sulfoxide r
43	69.5	7.8	168	2	S72898 hypotetrical prote
44	69.5	7.8	364	2	H87212 probable glucose e
45	69	7.7	266	2	S22511 chlorophyll a/b-bi

ALIGNMENTS

RESULT 1

A53019 collagen alpha 1(XVII) chain - human (fragment)

N:Contains: endostatin

C:Species: Homo sapiens (man)

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 31-Mar-2000

C/Accession: A53019

R/On: S.F.; Warman, M.L.; Seidlin, M.F.; Cheng, S.D.; Knoll, J.H.M.; Timmons, S.; Olsen, Genomics 19, 494-499, 1994

A/Title: Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localiz

A/Reference number: A53019; MWID:94245237; PMID:8188291

A/Accession: A53019

A/Molecule type: mRNA

A/Residues: 1-684 <OHA>

A/Cross-references: GB:L22548; NID:G348908; PIDN:AAAS1864.1; PID:G562794

A/Note: The cited accession number, L22548, is not in Genbank release 103

A/Note: In the authors' translation, 482-Gly is not shown, residues 483-490 are shifted

C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit

lated and subsequently O-glycosylated.

C/Comment: Different splice forms of collagen alpha 1(XVII) may be involved in perivas

C/Comment: Endostatin is released from collagen alpha 1(XVII) chain by the action of w

ay be useful in treating solid tumors.

C/Genetics:

A/Gene: GDB:COL18A1

A/Cross-references: GDB:138752; OMIM:120328

A/Map position: 21q22.3-21q22.3

C/Superfamily: unassigned collagenase

C/Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly

F/1-684/Product: collagen alpha 1(XVII) chain (fragment) #status predicted <MAT>

F/1-59/Domain: collagenous (fragment) #status predicted <CO4>

F/74-115/Domain: collagenous #status predicted <CO5>

F/129-201/Domain: collagenous #status predicted <CO6>

F/212-244/Domain: collagenous #status predicted <CO7>

F/257-278/Domain: collagenous #status predicted <CO8>

F/262-264/Region: cell attachment (R-G-D) motif

F/286-340/Domain: collagenous #status predicted <CO9>

F/354-371/Domain: collagenous #status predicted <CO10>

F/502-684/Product: endostatin #status predicted <KST>

F/509-684/Region: multiplexin collagen carboxyl-terminal similarity

Query Match

Best Local Similarity 100.0%; Score 893; DB 2; Length 684;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGRADPQFOOARAVGLAGTFRAFLSSRLDLYSVRRADRAVPIV 60

DB 514 VALNSPLSGMGRIGRADPQFOOARAVGLAGTFRAFLSSRLDLYSVRRADRAVPIV 573

QY 61 NLDELFLPSWEALFSGSGRLPLKGRARITSPDGKDVLRHPTWQKSVWMSDNGRRLTE 120

DB 574 NLDELFLPSWEALFSGSGRLPLKGRARITSPDGKDVLRHPTWQKSVWMSDNGRRLTE 633

Qy 121 SYCETWTEAPSATGQASLLGRLGQSAASCHHAYIVLCIENSFWTAS 170
 |||||
 Db 634 SYCETWTEAPSATGQASLLGRLGQSAASCHHAYIVLCIENSFWTAS 683

RESULT 2

B56101
 collagen alpha 1(XVIII) chain precursor, long splice form - mouse
 N/Contains: collagen alpha 1(XVIII) chain precursor, medium splice form; endostatin
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 31-Mar-2000
 A/Accession: B56101, C56101, S72450, S65595, P0675, A54072, A58816
 R/Rehm, M.; Philajanteml, T.
 J. Biol. Chem. 270, 4705-4711, 1995
 A/Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue homologous to rat and Drosophila frizzled proteins.
 A/Reference number: A56101, MUID:95181468, PMID:7876242
 A/Accession: B56101
 A/Molecule type: mRNA
 A/Residues: 1-562 <REH1>
 A/Cross-references: GB:U11637, NID:9618429, PIDN:AA52179.1, PID:9618430
 A/Experimental source: splice form clone PEL7.24
 A/Accession: C56101
 A/Molecule type: mRNA
 A/Residues: 1-239,487-562 <REH2>
 A/Cross-references: GB:U11637, NID:9618429
 A/Experimental source: splice form clones PB8.1, PB19, PE15.2
 R/Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 Submitted to the EMBL Data Library, August 1993
 A/Reference number: S72450
 A/Accession: S72450
 A/Molecule type: mRNA
 A/Residues: 487-1146, 'L', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'F', 1524-16
 A/Cross-references: EMBL:L22545, NID:9348968, PIDN:AA19787.1, PID:9511298
 R/Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
 A/Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa
 A/Reference number: A58370, MUID:94240111, PMID:8183893
 A/Accession: S65595
 A/Molecule type: mRNA
 A/Residues: 487-1512, 'L', 1514-1522, 'F', 1524-1683, 'V', 1685-1774 <OH2>
 A/Cross-references: EMBL:L22545
 R/Abe, N.; Muragaki, H.; Inoue, H.; Ninomiya, Y.
 Biochem. Biophys. Res. Commun. 196, 576-582, 1993
 A/Title: Identification of a novel collagen chain represented by extensive interruptions
 A/Reference number: P0675, MUID:94059075, PMID:8240330
 A/Accession: P0675
 A/Molecule type: mRNA
 A/Residues: 635-1774 <ABE>
 R/Rehm, M.; Hintikka, E.; Philajanteml, T.
 J. Biol. Chem. 269, 13929-13935, 1994
 A/Title: Primary structure of the alpha chain of mouse type XVIII collagen, partial structure
 A/Reference number: A54072, MUID:94245707, PMID:8188673
 A/Accession: A54072
 A/Molecule type: DNA; mRNA
 A/Residues: 1293-1403, 'R', 1405-1774 <REH3>
 A/Cross-references: GB:U03714, NID:9487733, PIDN:AAA20657.1, PID:9487734
 R/O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W.S.; Flynn, E.; Bi
 Cell 88, 277-285, 1997
 A/Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.
 A/Reference number: A58816, MUID:97160848, PMID:9008168
 A/Accession: A58816
 A/Molecule type: protein
 A/Residues: 1591-1610 <ORE>
 A/Experimental source: hemangioendothelium cells
 A/Note: Inhibits endothelial cell proliferation
 C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit labeled and subsequently O-glycosylated.
 C/Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in pericellular remodeling.
 C/Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un
 A/Genetics: ,

A/Genes: MGI:Coll18a1
 A/Cross-references: MGI:71175
 A/Map position: 10:41.0
 A/Intons: 1295/3, 1310/1, 1311/1, 1345/3, 1388/3, 1437/1, 1461/3, 1505/3, 1516/3, 1599/
 A/Note: the list of introns is incomplete
 C/Superfamily: unassigned collagens
 C/Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly
 F/1-1774/Product: collagen alpha 1(XVIII) chain precursor, long splice form #status pre
 F/1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, medium splice form #
 F/1-24/Dominant: signal sequence #status predicted <SIG>
 F/161-82/Region: frizzled similarity
 F/786-812/Dominant: collagenous #status predicted <CO01>
 F/823-896/Dominant: collagenous #status predicted <CO02>
 F/921-1042/Dominant: collagenous #status predicted <CO03>
 F/1066-1148/Dominant: collagenous #status predicted <CO04>
 F/1163-1204/Dominant: collagenous #status predicted <CO05>
 F/1218-1280/Dominant: collagenous #status predicted <CO06>
 F/1301-1333/Dominant: collagenous #status predicted <CO07>
 F/1346-1369/Dominant: collagenous #status predicted <CO08>
 F/1351-1353/Region: cell attachment (R-G-D) motif
 F/1377-1428/Dominant: collagenous #status predicted <CO09>
 F/1442-1459/Dominant: collagenous #status predicted <CO10>
 F/1591-1774/Product: endostatin #status predicted <EST>
 F/1598-1774/Region: multiplexin collagen carboxyl-terminal similarity
 F/354,361,947/Binding site: carbohydrate (asn) (covalent) #status predicted
 F/699,704,1716/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F/910,913,1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 87.1%; Score 778; DB 2; Length 1774;
 Best Local Similarity 85.8%; Pred. No. 1.3e-69;
 Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Qy 1 VALNSPLSGMGRGIRGADFOCFQPARAVGLAGTFRAFLSRLODLVSIVRADRAAPVIV 60
 |||||
 Db 1603 VALNPLSGMGRGIRGADFOCFQPARAVGLSGTFRAFLSRLODLVSIVRADRGSPVIV 1662
 Qy 61 NLKDELFPSEWALFSGSBEPKPGARIFSPDGKDVLRHPTTPQKSWHGSDPNGRLTE 120
 |||||
 Db 1663 NLKDELFPSEWALFSGSBEPKPGARIFSPDGKDVLRHPTTPQKSWHGSDPNGRLTE 1722
 Qy 121 SYCETWTEAPSATGQASLLGRLGQSAASCHHAYIVLCIENSFWTAS 169
 |||||
 Db 1723 SYCETWTEAPSATGQASLLGRLGQSAASCHHAYIVLCIENSFWTAS 1771
 RESULT 3
 A56101
 collagen alpha 1(XVIII) chain precursor, short splice form - mouse
 N/Contains: endostatin
 C/Species: Mus musculus (house mouse)
 C/Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 31-Mar-2000
 A/Accession: A56101, A58371, S72450, S65595
 R/Rehm, M.; Philajanteml, T.
 J. Biol. Chem. 270, 4705-4711, 1995
 A/Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue homologous to rat and Drosophila frizzled proteins.
 A/Reference number: A56101, MUID:95181468, PMID:7876242
 A/Accession: A56101
 A/Molecule type: mRNA
 A/Residues: 1-103 <REH1>
 A/Cross-references: GB:U11636, NID:9618427, PIDN:AA52178.1, PID:9618428
 R/Rehm, M.; Philajanteml, T.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994
 A/Title: Alpha1(XVIII), a collagen chain with frequent interruptions in the collagenous
 A/Reference number: A58371, MUID:94240112, PMID:8183894
 A/Accession: A58371
 A/Molecule type: mRNA
 A/Residues: 1-928 <REH2>
 A/Cross-references: GB:U11698, NID:9404754, PIDN:AAA7434.1, PID:9553894
 R/Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 Submitted to the EMBL Data Library, August 1993
 A/Reference number: S72450
 A/Accession: S72450

A:Molecule type: mRNA
A:Residues: 28-687, 'L', 689-734, 'F', 736-751, 'R', 753-1315 <OHM>
A:Cross-references: EMBL:L22545; NID:9348968; PIDN:AA19787.1; PID:9511298
R:Oh, S.P.; Kamaqata, Y.; Murgaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
A:Title: Isolation and sequencing of cDNAs for proteases with multiple domains of Gly-Xad
A:Reference number: A58370; MUID:94240111; PMID:8183893
A:Accession: 865595
A:Molecule type: mRNA
A:Residues: 28-1315 <OH8>
A:Cross-references: EMBL:L22545
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit labeled and subsequently O-glycosylated.
C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in peric
C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un
C:Comment: It is useful in treating solid tumors.
C:Genetics:
A:Gene: MGI:COL18A1
A:Cross-references: MGI:71175
A:Map position: 10:41.0
C:Superfamily: unaassigned collagens
C:Keyword: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc
F:1-25/Domain: signal sequence #status predicted <Sig>
F:24-235/Region: thrombospondin amino-terminal similarity
F:26-1315/Product: collagen alpha 1(XVIII) chain, short splice form #status predicted <M
F:327-353/Domain: collagenous #status predicted <CO1>
F:364-437/Domain: collagenous #status predicted <CO2>
F:462-583/Domain: collagenous #status predicted <CO3>
F:607-689/Domain: collagenous #status predicted <CO4>
F:704-745/Domain: collagenous #status predicted <CO5>
F:759-831/Domain: collagenous #status predicted <CO6>
F:842-874/Domain: collagenous #status predicted <CO7>
F:887-910/Domain: collagenous #status predicted <CO8>
F:892-894/Region: cell attachment (R-G-D) motif
F:918-969/Domain: collagenous #status predicted <CO9>
F:983-1000/Domain: collagenous #status predicted <CO10>
F:1132-1315/Product: endostatin #status predicted <EST>
F:1139-1315/Region: multiplexin collagen carboxyl-terminal similarity
F:126/488/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:172-228/Diulfide bonds: #status predicted
F:240.245.1257/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:451.454.594/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 86.8%; Score 775; DB 2; Length 1315;
Best Local Similarity 85.2%; Pred. No. 1.8e-69;
Matches 144; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTRAFLLSLRLDLYSIVRRADRAVPYV 60
DB 1144 VALNTPLSGMRGIRGADFOCFQOARAVGLAGTRAFLLSLRLDLYSIVRRADRAVPYV 1203

QY 61 NLKDELLFSPSEALFSGSEGLKRGARIFSPDGKDVLRHPWPCKSVWHSDDPGRRLTE 120
DB 1204 NLKDEVLSPWMDISFSSQGVOPGARIFSPDGKDVLRHPWPCKSVWHSDDPGRRLTE 1263

QY 121 SYCETWRTAPSATGQASLLGRLGQSAASCHNAVIVLCIENSFMTA 169
DB 1264 SYCETWRTETTGATGQASLLGRLGQSAASCHNAVIVLCIENSFMTS 1312

RESULT 4
A53317
collagen alpha 1(XV) chain precursor - human
N:Alternate names: procollagen alpha 1(XV) chain
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 31-Mar-2000
A:Accession: A53317; A53146; S28778
R:Kivirikko, S.; Heineken, P.; Rehn, M.; Honkanen, N.; Myers, J.C.; Pihlajaniemi, T.
J. Biol. Chem. 269, 4773-4779, 1994
A:Title: Primary structure of the alpha chain of human type XV collagen and exon-intron
A:Reference number: A53317; MUID:94148920; PMID:8106446
A:Accession: A53317
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-1388 <KIV>
A:Cross-references: GB:L25280
A:Note: nucleotide sequence and conceptual translation not complete
R:Murgaki, Y.; Abe, N.; Nimmiya, Y.; Olsen, B.R.; Ooshima, A.
J. Biol. Chem. 269, 4042-4046, 1994
A:Title: The human alpha1(XV) collagen chain contains a large amino-terminal non-triple
A:Reference number: A53146; MUID:94140817; PMID:8307960
A:Accession: A53146
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-9, 'S', 11-48, 'V', 50-94, 'A', 96-145, 'A', 151-203, 'V', 205-408, 'A', 410-569 <MUR
A:Cross-references: GB:D21230; NID:9415605; PIDN:BA04762.1; PID:di005294; PID:9460703
R:Myers, J.C.; Kivirikko, S.; Gordon, M.K.; Dion, A.S.; Pihlajaniemi, T.
Proc. Natl. Acad. Sci. U.S.A. 89, 10144-10148, 1992
A:Title: Identification of a previously unknown human collagen chain, alpha1(XV), charac
A:Reference number: 628778; MUID:93066196; PMID:1279671
A:Accession: S28778
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 544-640, 'P', 642-811, 'P', 813-1252 <MYB>
C:Genetics:
A:Gene: GDB:COL5A1
A:Cross-references: GDB:132578; OMIM:120325
A:Map position: 9q21-9q22
C:Superfamily: unaassigned collagens
F:1-22/Domain: signal sequence #status predicted <Sig>
F:23-1388/Product: collagen alpha 1(XV) chain #status predicted <MAT>
F:1216-1388/Region: multiplexin collagen carboxyl-terminal similarity

Query Match 56.4%; Score 504; DB 2; Length 1388;
Best Local Similarity 56.9%; Pred. No. 3.4e-42;
Matches 95; Conservative 27; Mismatches 41; Indels 4; Gaps 1;

QY 2 ALNSPLSGMGRGIRGADFOCFQOARAVGLAGTRAFLLSLRLDLYSIVRRADRAVPYV 61
DB 1222 ALNMPSPSGDIR---ADFQCFQARAGLSTYRAFLSLRLDLYSIVRAERYSLPIVN 1277

QY 62 LKDELLFSPSEALFSGSEGLKRGARIFSPDGKDVLRHPWPCKSVWHSDDPGRRLTES 121
DB 1278 LKQVLEFNMWDSIFSGHGGFNMWITPISFDGKDVLRHPWPCKSVWHSDDPGRRLVDN 1337

QY 122 YCETWRTAPSATGQASLLGRLGQSAASCHNAVIVLCIENSFMT 168
DB 1338 YCETWRTADRAVGLAGLSTGKILQKAVCANRLIVLCIENSFMT 1384

RESULT 5
T22002
hypothetical protein F39H11.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A:Accession: T22002
R:White, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19500
A:Accession: T22002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-650 <WIL>
A:Cross-references: EMBL:281079; PIDN:CA03084.1; GSPDB:GN00019; CESP:F39H11.4
C:Genetics:
A:Gene: CESP:F39H11.4
A:Map position: 1
A:Introns: 109/3; 154/1; 357/1; 420/3; 464/3; 566/2; 594/1; 628/3

Query Match 40.0%; Score 357; DB 2; Length 650;
Best Local Similarity 44.0%; Pred. No. 8.4e-28;
Matches 73; Conservative 24; Mismatches 63; Indels 6; Gaps 4;

QY 1 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTRAFLLSLRLDLYSIVRRADRAVPYV 60
DB 1144 VALNTPLSGMRGIRGADFOCFQOARAVGLAGTRAFLLSLRLDLYSIVRRADRAVPYV 1203

Db 140 LPPNGNDEEDLPKKDDIRIDKPEBQWMAEDSEGRGMIPVYVEKRR-----PASA 195
 QY 136 QASSLGGR 144
 Db 196 SVSALIGR 204

RESULT 10
 JMW0047
 class I cytokinase receptor precursor - human
 N:Alternate names: MSX-1
 C:Species: Homo sapiens (man)
 C>Date: 17-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
 C:Accession: JMW0047
 R:Spiecher, C.A.; Grant, F.J.; Baumgartner, J.W.; Preenell, S.R.; Schrader, S.K.; Yamagi
 Biochem. Biophys. Res. Commun. 246, 82-90, 1998
 A>Title: Cloning and characterization of a novel class I cytokine receptor.
 A:Reference number: JMW0047; MUID:98262921; PMID:9600072
 A:Accession: JMW0047
 A:Molecule type: mRNA
 A:Residues: 1-636 <SPR>
 A:Cross-references: GB:AF053004; NID:G3153240; PIDN:AAC39755.1; PID:G3153241
 A:Experimental source: brain
 C:Genetics:

A:Map position: 19p13.11
 C:Keywords: glycoprotein
 P:1-33/Domain: signal sequence #status predicted <SIG>
 P:554-540/Domain: transmembrane #status predicted <TM>
 P:554-561/Domain: cytoplasmic #status predicted <CTP>
 P:51,76,302,311,374,382,467/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 8.5%; Score 76; DB 2; Length 636;
 Best Local Similarity 24.2%; Pred. No. 15;
 Matches 39; Conservative 12; Mismatches 46; Indels 64; Gaps 7;
 QY 11 MRGIRGADF-----QCRQARAVGLACTFPAFLSSRLQDL----- 45
 Db 1 MRGRGAPFWLWPLPKLALPLWLVLFQRTRPQGSAGPLCCYGVGPDILNCSWEPLGDL 60
 QY 46 -----YSIVRADRAVPI-----VNLDELFF-----PSWEALFS 76
 Db 61 GASELHLGOSKTRSNKTQVAVAAGRSWAIPREBQITMSBDKLLVMGTAKGAPLMPVFFV 120

QY 77 GSEGPLKPGA-RI---FSFDGKDVLR-----HPTWPKSV 107
 Db 121 NLETKMKPNAPRLGPVDVDFSEDDPLEATVMAPTWPSHV 161

RESULT 11
 T04377
 probable pullulanase (EC 3.2.1.41) - barley
 N:Alternate names: pullulanase
 C:Species: Hordeum vulgare (barley)
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
 C:Accession: T04377
 R:Jok, F.; Kristensen, M.; Plancho, V.; Leah, R.; Svendsen, I.; Svenson, B.
 submitted to the EMBL Data Library, December 1997
 A:Description: Isolation and characterization of starch debranching enzyme, limit dextri
 A:Reference number: Z15320
 A:Accession: T04377
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-904 <LOK>
 A:Cross-references: EMBL:AF022725; NID:92502057; PIDN:AMD04189.1; PID:92677837
 A:Experimental source: cv. Igri
 C:Genetics:
 A:Gene: HvLUD99
 A:Introns: 21/3; 70/1; 87/3; 124/1; 169/3; 235/2; 285/2; 305/3; 346/1; 376/2; 407/3; 439
 C:Keywords: glycosidase; hydrolase

Query Match 8.5%; Score 75.5; DB 2; Length 904;
 Best Local Similarity 22.5%; Pred. No. 26;
 Matches 45; Conservative 19; Mismatches 67; Indels 69; Gaps 8;

QY 5 SPLSGMRGIRGADFQCFQOARAVGLAGT-----FRAFLSSRLQDLVYSIVR----- 50
 Db 62 SPENG---GIQGYDSKVELQPSAGLPETVTKFPISSYRAKVPSSVDVSLVNCQLV 118
 QY 51 ----RADRAVPIVNI-----KDELLPSSWE-----ALPSG 77
 Db 119 VASFGADGKHVDVTGLQPLGVLDMPAYTPGLGAVSESDSVSLHMAPTAGVSVCFPDG 178
 QY 78 SEGR-----LKPGRIFSPDGK-----DVLHPTWPKSVHMGSDPNGRRLT 119
 Db 179 PAFPALETVOLKESNGVSWVTGPREMENRYLYEVGVV-HPTAQLVKCLAGDPARSIS 237
 QY 120 ESYCEFTWTEAPSATQOASS 139
 Db 238 ANGARTWLDINNETHKPPAS 257

RESULT 12
 B89781
 conserved hypothetical protein SA0184 [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: B89781
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu
 ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: B89758; MUID:21311952; PMID:11418146
 A:Accession: B89781
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-351 <KUR>
 A:Cross-references: GB:BA000018; PID:G13700106; PIDN:BA041405.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA0184

Query Match 8.4%; Score 75; DB 2; Length 351;
 Best Local Similarity 23.3%; Pred. No. 9.2;
 Matches 42; Conservative 17; Mismatches 67; Indels 54; Gaps 8;
 QY 11 MRGIRGADFQCFQOARAVGLAGTFRFLSSRLQDL-----YSIVRADRA-AVPIY 60
 Db 97 IEALMAQGLKCLMAIIS-----RELITSHQQLANDFTLISFCNRYRRPDTGLSDLV 151
 QY 61 NLKDELLFSPWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 120
 Db 152 NKNKELIY-----QFNPKAQIYGFIVGSLRGPL-----HKGLPT----- 186
 QY 121 SYCEFTWTEAP-----SATQOASSLLGRLLGQSAAS-----CHNAVIVLCIENSFMT 168
 Db 187 -LEATRSHHPVAAKLQETGVSEVLVDGSLIEMKQAKQLIDFCGRHFTLLICEBVFDT 244

RESULT 13
 S56015
 gastric mucin MUC5AC - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Apr-2000
 C:Accession: S56015; S53361
 R:Klomp, L.W.J.; van Rens, L.; Strouds, G.J.
 Biochem. J. 308, 831-838, 1995
 A>Title: Cloning and analysis of human gastric mucin cDNA reveals two types of conserved
 A:Reference number: S56015; MUID:97104281; PMID:8948439
 A:Accession: S56015
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-850 <KLO>
 A:Cross-references: EMBL:X81649; NID:9547516; PIDN:CA547309.1; PID:9547517
 R:Guyonnet-Duperat, V.; Audie, J.P.; Debaillevet, V.; Laine, A.; Bulsine, M.P.; Gallegue-
 Biochem. J. 305, 211-219, 1995

A>Title: Characterization of the human mucin gene MUC5AC: a consensus cysteine-rich dom
A/Reference number: S53361; MUID:95126907; PMID:7826332
A/Accession: S53361
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 648-678, 'L', 680-733, 'L', 735-760 <GV>
A/Cross-references: EMBL:Z34280; NID:9563380; PID:9563381
A/Experimental source: clone JUL32
A/Note: this publication is not cited in GenBank entry HSMUCIN5, release 113.0

Query Match 8.2%; Score 74.5; DB 2; Length 850;
Best Local Similarity 24.4%; Pred. No. 30;

Matches 39; Conservative 23; Mismatches 71; Indels 25; Gaps 6;

QY 13 GIRADFOCFQOARAVAGTFRALFSLRLQDLVSVRRADR-AAVPIVNLKDELLFSPSW 71
DB 592 GINGDEPTFQNLDEGT--TF-----CESPRSVQCRASFPNTPLADGQVIGCHT 642
QY 72 EALFSGSEGPLK-----GARIFSDGKDLRHPTWPKSVW-----HGSDPGRRLTE 120
DB 643 EGLICLNKNGLPICVYVEIRIQCCETVNVCRDITRPKVTATRPHPRTGAQTQTTFT 702
QY 121 SYCETWRTAPASATGQ---ASSLIGRLGQSAASCH 154
DB 703 THMPSASTEPTATSRGPTATSVTGTHTTPVTRNCH 740

RESULT 14

B91052
hypothetical protein EC63386 [imported] - Escherichia coli (strain O157:H7, substrain R1
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C/Accession: B91052
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gaeswara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Ref. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A/Reference number: A9629; MUID:21156231; PMID:11258796
A/Accession: B91052
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1653 <HAY>
A/Cross-references: GB:BA000007; PID:BA836809.1; PID:g13362856; GSPDB:GN00154
A/Experimental source: strain O157:H7, substrain R1MD 050952
C/Genetics:
A/Gene: EC63386

Query Match 8.2%; Score 73.5; DB 2; Length 1653;

Best Local Similarity 24.4%; Pred. No. 84;
Matches 32; Conservative 21; Mismatches 47; Indels 31; Gaps 5;

QY 29 GLAGTFRAPFLSSRLQDLY-----SIVRRADRAAVPIVNLKDELLFSPWEALFSGSE 79
DB 380 GAFYSKQFMFGPRDLYRGETVILNGLRDADGKALPNQPIKLDVIKPDQVLSVVS 439
QY 80 GPLKPGARIFSDGKDLRHPTWPKS-----VNH---GSDPGRRLTESYCEWTRTE-- 129
DB 440 QP-----ENGLYHFTWPLDSNATGMHIRANTGDNQYRWMDPHVEDFMPERM 487
QY 130 APSATGQASSL 140
DB 488 ALNLTGKPTPL 498

RESULT 15

F85896
hypothetical protein Z3787 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: F85896
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.O.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: F85896
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1653 <STO>
A/Cross-references: GB:AE005174; NID:g12516921; PID:AA057634.1; GSPDB:GN00145; UNCF:23
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: Z3787

Query Match 8.2%; Score 73.5; DB 2; Length 1653;

Best Local Similarity 24.4%; Pred. No. 84;
Matches 32; Conservative 21; Mismatches 47; Indels 31; Gaps 5;

QY 29 GLAGTFRAPFLSSRLQDLY-----SIVRRADRAAVPIVNLKDELLFSPWEALFSGSE 79
DB 380 GAFYSKQFMFGPRDLYRGETVILNGLRDADGKALPNQPIKLDVIKPDQVLSVVS 439
QY 80 GPLKPGARIFSDGKDLRHPTWPKS-----VNH---GSDPGRRLTESYCEWTRTE-- 129
DB 440 QP-----ENGLYHFTWPLDSNATGMHIRANTGDNQYRWMDPHVEDFMPERM 487
QY 130 APSATGQASSL 140
DB 488 ALNLTGKPTPL 498

Search completed: February 11, 2003, 20:43:09
Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:36:43 ; Search time 36 Seconds
(without alignments)
629.239 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893
Sequence: 1 VALNSPLSGMKRGIRGADFDQ.....ASCHNAVYLCIENSFMTAS 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	178	21	AA94324
2	893	100.0	178	22	AAU0930
3	893	100.0	179	22	AAU0901
4	893	100.0	182	21	AA82839
5	893	100.0	182	21	AA94323
6	893	100.0	182	21	AA95962
7	893	100.0	182	22	AAU0897
8	893	100.0	182	23	AAU77951
9	893	100.0	183	20	AAU08693
10	893	100.0	183	20	AAU02113

11	893	100.0	183	21	AA830493
12	893	100.0	183	21	AA816451
13	893	100.0	183	21	AA90771
14	893	100.0	183	21	AA970252
15	893	100.0	183	22	AAU0896
16	893	100.0	183	22	AA849379
17	893	100.0	183	23	AAU97132
18	893	100.0	183	23	AAU49503
19	893	100.0	183	23	AAU48895
20	893	100.0	193	23	AAU76690
21	893	100.0	195	21	AAW90874
22	893	100.0	216	21	AA830495
23	893	100.0	275	23	AAU76689
24	893	100.0	310	23	AAU76688
25	893	100.0	682	23	ABP41878
26	893	100.0	684	18	AAW26327
27	893	100.0	684	20	AAU25113
28	893	100.0	684	23	AAU71357
29	893	100.0	1301	20	AAW92296
30	893	100.0	1336	20	AAU08694
31	893	100.0	1516	22	ABB83471
32	889	99.6	181	22	AAU00898
33	889	99.6	183	22	AA849810
34	886	99.2	184	22	AAU76717
35	885	99.1	180	22	AAU00899
36	885	99.1	193	21	AAW90877
37	882	98.8	271	21	AA808407
38	783	87.7	184	21	AAV70265
39	783	87.7	184	23	AAU17430
40	783	87.7	230	23	AAU17429
41	779	87.2	684	20	AAV25114
42	778	87.1	183	23	AAU49504
43	778	87.1	191	22	AA849380
44	778	87.1	191	21	AA828398
45	778	87.1	191	23	AAU77950

ALIGNMENTS

RESULT 1
AA94324
ID AA94324 standard; Protein; 178 AA.
XX
AC AA94324;
XX
DT 11-AUG-2000 (first entry)
XX
DE Alternate human endostatin protein.
XX
KW Human; endothelial cell proliferation inhibitor; collagen XVIII;
KW angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
KW vasotrophic; dermatological; ophthalmological; vulnary;
KW antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;
KW ocular angiogenic disease; atherosclerosis; scleroderma;
KW myocardial angiogenesis; telangiectasia; angiofibroma;
KW wound granulation.
XX
OS Homo sapiens.
XX
PN WO200026368-A2.
XX
PD 11-MAY-2000.
XX
PF 01-NOV-1999; 99WO-US25605.
XX
PR 30-OCT-1998; 98US-0106343.
XX
PR 20-MAY-1999; 99US-0315689.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI O'Reilly MS; Folkman MJ;
XX

Amino acid sequenc
Human endostatin p
Human angiogenesis
Human angiogenesis
Human endostatin(T
Human endostatin S
Human endostatin.
Human endostatin p
Human endostatin p
Synthetic plasmid
Human HMW endostac
Amino acid sequenc
Synthetic plasmid
Human collagen XVI
Human collagen XVI
Human ovarian anti
Human alpha-1 coll
Human alpha (XVII
Human collagen typ
Human alpha-1 (XVI
Human collagen 18
Human collagen XVI
Human endostatin(T
Human endostatin p
Human vascular end
Human endostatin(T
Human HMW endostac
A human angiogenes
Canine angiogenesi
Canine endostatin.
Canine pro-endosta
Mouse alpha (XVII
Human endostatin p
Murine endostatin
Murine endostatin
Amino acid sequenc

DR WPI; 2000-365617/31.
XX N-PSDB; AAA27005.
PT Novel endostatin capable of inhibiting endothelial cell proliferation
PT and angiogenesis, useful for treating angiogenesis-dependent cancers
PT and as birth control agents -
XX
XX Claim 3; Page 39; 68pp; English.
XX
CC The present sequence is an alternate functional endostatin
CC protein. When the human endostatin gene sequence AAA27004 is
CC recombinantly expressed, an observable doublet of protein results, both
CC versions of which are functional endostatin proteins. The present
CC endostatin variant is the same as the protein encoded by AAA27004 minus
CC the first four amino acids. Recombinant mouse endostatin (20 mg/kg) was
CC administered subcutaneously to mice implanted with Lewis lung carcinomas.
CC There was tumour mass regression non-detectable levels after 12 days of
CC therapy due to the angiogenesis inhibitory activity of endostatin. Thus
CC the protein is useful for treatment of angiogenesis- dependent cancers.
CC The polynucleotide and polypeptide sequences of this endostatin are
CC useful for treating and diagnosis of tumours, ocular angiogenic
CC diseases, Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma
CC and wound granulation, for treatment of diseases related to excessive or
CC abnormal stimulation of endothelial cells e.g. intestinal adhesions,
CC atherosclerosis, scleroderma. The protein may also be useful as a birth
CC control agent by reducing or preventing uterine vascularisation. The
CC gene for endostatin may be isolated from cells or tissue that express
CC high levels of endostatin, eg. tumour cells, by generating cDNA from
CC mRNA using reverse transcriptase and then amplifying the DNA sequence.
XX
SQ Sequence 178 AA;
Query Match 100.0%; Score 893; DB 21; Length 178;
Best Local Similarity 100.0%; Pred. No.3.5e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTRAFLSRLDLYSIVRRADRAAVPIV 60
DB 9 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTRAFLSRLDLYSIVRRADRAAVPIV 68
QY 61 NLDELFLFPMWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSQDPNGRRLTE 120
DB 69 NLKELLFPMWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSQDPNGRRLTE 128
QY 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 129 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 178
RESULT 2
AAU00900
ID AAU00900 standard; Protein; 178 AA.
XX
AC AAU00900;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human Endostatin(TM) N-terminal deletion mutant protein#2.
XX
XX Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
XX haemangioendothelioma; leukaemia; tumour metastasis; benign tumour;
XX haemangioma; acoustic neuroma; neurofibroma; trachoma; rubosis;
XX pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
XX ocular angiogenic disease; diabetic retinopathy; macular degeneration;
XX retinopathy of prematurity; macular corneal graft rejection;
XX neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
XX myocardial angiogenesis; plaque neovascularisation; telangiectasia;
XX haemophilic joint; angiodiroma; wound granulation; variant;
XX mutant; mutein.
XX
XX Homo sapiens.
OS
XX

PN WO200119989-A2.
XX
XX 22-MAR-2001.
PD
XX
XX 14-SEP-2000; 2000WO-US25166.
PP
XX
XX 14-SEP-1999; 99US-0153698.
PR
XX
XX (ENTR-) ENTREMED INC.
PA
XX
XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
PI Bermejo LL, Mistry FR, Shepard SR, Schrimmer JL;
PI N-PSDB; AAS00868.
DR
XX
XX WPI; 2001-244802/25.
DR N-PSDB; AAS00868.
PT Producing Endostatin protein for treating angiogenesis mediated
PT diseases such as solid tumours, comprises recombinantly producing the
PT protein using an expression system, and recovering and purifying the
PT protein -
XX
XX
PS Claim 5; Page 33; 67pp; English.
XX
XX The sequence represents Human Endostatin(TM) N-terminal deletion mutant
XX protein lacking the N-terminal 4 amino acids and the C-terminal lysine,
XX a natural variant recovered from fermentations of Pichia pastoris
XX cultures harbouring an expression plasmid containing the Endostatin(TM)
XX DNA sequence given in AAS00868. The new method of the invention is
XX useful for producing, recovering and purifying Endostatin(TM) from
XX biological sources, such as biological fluids, tissues, cells, culture
XX media, and fermentation media. Endostatin(TM) is useful for treating
XX angiogenesis mediated diseases such as solid tumours, blood borne
XX tumours, leukaemias, tumour metastases, benign tumours, e.g. haemangioma,
XX acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas,
XX rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
XX e.g., diabetic retinopathy, retinopathy of prematurity, macular
XX degeneration, corneal graft rejection, neovascular glaucoma, colon
XX cancer, retrolental fibroplasia, rubosis, Osler-Webber Syndrome,
XX myocardial angiogenesis, plaque neovascularisation, telangiectasia,
XX haemophilic joints, angiodiroma, and wound granulation. Endostatin(TM)
XX is also useful for treating disease of excessive or abnormal stimulation
XX of endothelial cells such as intestinal adhesions, atherosclerosis,
XX scleroderma and hypertrophic scars. Higher yields of more purified, and
XX biologically active Endostatin(TM) are obtained by the new method.
XX Endostatin(TM) can be stored in buffers for extended periods of time, and
XX also subjected to lyophilisation, while preserving biological activity.
XX Centrifugation of broth from fermentation steps in production is avoided,
XX preventing unwanted potential cellular lysis and contamination with
XX additional proteins, pigments, enzymes and other cellular chemicals and
XX debris.
XX
SQ Sequence 178 AA;
Query Match 100.0%; Score 893; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No.3.5e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTRAFLSRLDLYSIVRRADRAAVPIV 60
DB 9 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTRAFLSRLDLYSIVRRADRAAVPIV 68
QY 61 NLDELFLFPMWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSQDPNGRRLTE 120
DB 69 NLDELFLFPMWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSQDPNGRRLTE 128
QY 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 129 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 178
RESULT 3
AAU00901
ID AAU00901 standard; Protein; 179 AA.


```

XX AC AAU0901;
XX DT 04-JUL-2001 (first entry)
XX DE Human Endostatin(TM) N-terminal mutant protein#1.
XX
XX KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;
XX KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
XX KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;
XX KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
XX KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
XX KW retinopathy of prematurity; macular corneal graft rejection;
XX KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
XX KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
XX KW haemophilic joint; angiofibroma; wound granulation; mutant; mutein.
XX
XX OS Homo sapiens.
XX PN WO200119989-A2.
XX
XX PD 22-MAR-2001.
XX
XX PF 14-SEP-2000; 2000WO-US25166.
XX PR 14-SEP-1999; 99US-0153698.
XX
XX PA (ENTR-) ENTREMED INC.
XX PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;
XX PI Bernejo LT, Mistry FR, Shepard SR, Schrimmer JI;
XX DR N-PSDB; AAS00868.
XX
XX PT Producing Endostatin protein for treating angiogenesis mediated
XX PT diseases such as solid tumours, comprises recombinantly producing the
XX PT protein using an expression system, and recovering and purifying the
XX PT protein.
XX
XX PS Claim 5; Page 32; 67pp; English.
XX
XX CC The sequence represents a Human Endostatin(TM) N-terminal deletion
XX CC mutant lacking the N-terminal 4 amino acids. The new method of the
XX CC invention is useful for producing, recovering and purifying Endostatin
XX CC (TM) from biological sources, such as biological fluids, tissues, cells,
XX CC culture media, and fermentation media. Endostatin(TM) is useful for
XX CC treating angiogenesis mediated diseases such as solid tumours, blood
XX CC borne tumours, leukaemia, tumour metastases, benign tumours, e.g.
XX CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
XX CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
XX CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
XX CC degeneration, corneal graft rejection, neovascular glaucoma, colon
XX CC cancer, retrolental fibroplasia, rubecosis, Osler-Webber Syndrome,
XX CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
XX CC haemophilic joint, angiofibroma, and wound granulation. Endostatin(TM)
XX CC is also useful for treating disease of excessive or abnormal stimulation
XX CC of endothelial cells such as intestinal adhesions, atherosclerosis,
XX CC scleroderma and hypertrophic scars. Higher yields of more purified, and
XX CC biologically active Endostatin(TM) are obtained by the new method.
XX CC Endostatin(TM) can be stored in buffers for extended periods of time, and
XX CC also subjected to lyophilisation, while preserving biological activity.
XX CC Centrifugation of broth from fermentation steps in production is avoided,
XX CC preventing unwanted potential cellular lysate and contamination with
XX CC additional proteins, pigments, enzymes and other cellular chemicals and
XX CC debris.
XX
XX SQ Sequence 179 AA;

```

```

QY 1 VALNSPLSGMGRIGRADPQCFOQARAAGLAGTFRAFLSSRLQDLYSIVRRADRAVPVY 60
DB 9 VALNSPLSGMGRIGRADPQCFOQARAAGLAGTFRAFLSSRLQDLYSIVRRADRAVPVY 68
QY 61 NLKDELLPFSWELPFGSGRPLKPGARIFSPGKQVLRHPTVPQKSVHMGSDPNGRRLTE 120
DB 69 NLKDELLPFSWELPFGSGRPLKPGARIFSPGKQVLRHPTVPQKSVHMGSDPNGRRLTE 128
QY 121 SYCETWRTAPASATGQASSLLGGRLLGQSAACHAYIVLCIENSFMTAS 170
DB 129 SYCETWRTAPASATGQASSLLGGRLLGQSAACHAYIVLCIENSFMTAS 178

RESULT 4
ID AAB28399 standard; Protein; 182 AA.
XX
XX AC AAB28399;
XX DT 19-FEB-2001 (first entry)
XX
XX DE Human endostatin.
XX
XX KW Human; endostatin; cytostatic; antiproliferative;
XX KW vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;
XX KW cancer; vascularised solid tumour.
XX
XX OS Homo sapiens.
XX PN WO200064946-A2.
XX PD 02-NOV-2000.
XX
XX PF 28-APR-2000; 2000WO-US11367.
XX PR 28-APR-1999; 99US-0131432.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX
XX PI Thorpe BE, Brekken RA;
XX DR WPI; 2000-687317/67.
XX
XX PT Immunogenic composition for the treatment and diagnosis of cancer
XX PT comprises an anti-VEGF (vascular endothelial growth factor) antibody
XX PT binding the same epitope as the monoclonal antibody ATCC PTA 1595 -
XX
XX PS Example 10; Page 291-292; 298pp; English.
XX
XX CC The present invention relates to anti-Vascular Endothelial Growth Factor
XX CC (VEGF) antibodies that bind to the same epitope as the monoclonal
XX CC antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to
XX CC the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF
XX CC receptor VEGFR1. The present sequence is human endostatin. Endostatin
XX CC may be conjugated onto the anti-VEGF antibodies of the present invention.
XX CC The anti-VEGF antibodies of the present invention are useful for the
XX CC treatment and diagnosis of cancer, especially vascularised solid tumours.
XX
XX SQ Sequence 182 AA;

```

```

Query Match 100.0%; Score 893; DB 21; Length 182;
Best Local Similarity 100.0%; Pred. No. 3,66-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VALNSPLSGMGRIGRADPQCFOQARAAGLAGTFRAFLSSRLQDLYSIVRRADRAVPVY 60
DB 13 VALNSPLSGMGRIGRADPQCFOQARAAGLAGTFRAFLSSRLQDLYSIVRRADRAVPVY 72
QY 61 NLKDELLPFSWELPFGSGRPLKPGARIFSPGKQVLRHPTVPQKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLPFSWELPFGSGRPLKPGARIFSPGKQVLRHPTVPQKSVHMGSDPNGRRLTE 132
QY 121 SYCETWRTAPASATGQASSLLGGRLLGQSAACHAYIVLCIENSFMTAS 170

```

Db 133 SYCETWRTPEASATGQASSILGRLIGQSAASCHHAYIVLCIENSFWMTAS 182

RESULT 5
AAV94323
ID AAV94323 standard; Protein; 182 AA.

AC AAY94323;

DT 11-AUG-2000 (first entry)

DE Human endostatin protein.

Human; endothelial cell proliferation inhibitor; collagen XVIII;
angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
vasotropic; dermatological; ophthalmological; vulvar; anti-
arteriosclerotic; antidiabetic; haemostatic; contraceptive;
ocular angiogenic disease; atherosclerosis; scleroderma;
myocardial angiogenesis; telangiectasia; angiodysplasia;
wound granulation.

OS Homo sapiens.

PN WO200026368-A2.

PD 11-MAY-2000.

PF 01-NOV-1999; 99WO-US25605.

PR 30-OCT-1998; 98US-0106343.

PR 20-MAY-1999; 99US-0315689.

XX (CHIL-) CHILDRENS MEDICAL CENT.

PI O'Reilly MS, Folkman MJ;

DR WPI; 2000-365617/31.

DR N-PSDB; AAA27004.

PT Novel endostatin capable of inhibiting endothelial cell proliferation
and angiogenesis, useful for treating angiogenesis-dependent cancers
and as birth control agents -
PS Claim 2; Page 38; 68pp; English.

CC The present sequence is an endostatin protein which is the carboxy
terminal protein of human collagen XVIII. Recombinant mouse endostatin
(20 mg/kg) was administered subcutaneously to mice implanted with Lewis
lung carcinomas. There was tumour mass regression non-detectable levels
after 12 days of therapy due to the angiogenesis inhibitory activity of
endostatin. Thus the protein is useful for treatment of angiogenesis-
dependent cancers. The polynucleotide and polypeptide sequences of this
endostatin are useful for treating and diagnosis of tumours, ocular
angiogenic diseases, Osler-Webber syndrome, myocardial angiogenesis,
CC plaque neovascularisation, telangiectasia, haemophilic joints,
CC angiodysplasia and wound granulation, for treatment of diseases related to
excessive or abnormal stimulation of endothelial cells e.g. intestinal
adhesions, atherosclerosis, scleroderma. The protein may also be useful
as a birth control agent by reducing or preventing uterine
vascularisation. The gene for endostatin may be isolated from cells or
tissue that express high levels of endostatin, eg. tumour cells, by
CC generating cDNA from mRNA using reverse transcriptase and then amplifying
the DNA sequence.

XX Sequence 182 AA;

Query Match 100.0%; Score 893; DB 21; Length 182;

Best Local Similarity 100.0%; Pred. No. 3.6e-101; Mismatches 0; Gaps 0;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSRLQDLYSVRADRAVPIV 60

Db 13 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSRLQDLYSVRADRAVPIV 72

QY 61 NLKDELLFPMSEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 120

Db 73 NLKDELLFPMSEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 132

QY 121 SYCETWRTPEASATGQASSILGRLIGQSAASCHHAYIVLCIENSFWMTAS 170

Db 133 SYCETWRTPEASATGQASSILGRLIGQSAASCHHAYIVLCIENSFWMTAS 182

RESULT 6
AAV59622
ID AAV59622 standard; protein; 182 AA.

AC AAY59622;

DT 14-MAR-2000 (first entry)

DE Human endostatin protein fragment.

Human; scatter factor activity; human; tubulogenesis; psoriasis;
metastatic cancer; tumorigenesis; ocular angiogenic disease;
rheumatoid arthritis; Osler-Webber syndrome; telangiectasia;
haemophilic joint; angiodysplasia; wound granulation.

OS Homo sapiens.

PN WO962944-A2.

PD 09-DEC-1999.

PF 03-JUN-1999; 99WO-US12278.

PR 03-JUN-1998; 98US-0087890.

PR 10-JUL-1998; 98US-0092393.

PR 01-SEP-1998; 98US-0098790.

XX (CHIL-) CHILDRENS MEDICAL CENT.

PI Javaherian K, Folkman MJ;

DR WPI; 2000-072833/06.

PT New endostatin oligomers, used for treating e.g. tumours -
PS Disclosure; Page 6; 44pp; English.

CC This sequence is a fragment of the human endostatin protein. Endostatin
is an approximately 20kD C-terminal globular domain of the collagen-like
protein collagen XVIII. Protein oligomers consisting of more than one
endostatin monomer have anti-tubulogenic effects and induce
reorganization of the actin cytoskeleton. The oligomer has scatter factor
activity. The oligomers induce the destruction of tubular lumens and
elongation of cells, and inhibit tubulogenesis and tumorigenesis. The
oligomers can also be used to treat metastatic cancers, tumours,
CC rheumatoid arthritis, psoriasis, ocular angiogenic disease, Osler-Webber
syndrome, plaque neovascularisation, telangiectasia, haemophilic
joints, angiodysplasia and wound granulation. The oligomers can also be
used to treat diseases that have angiogenesis as a pathological
consequence e.g. ulcers. The endostatin oligomers can also be used to
develop affinity columns for isolating antibodies or receptors. Passive
antibody therapy using antibodies that specifically bind endostatin
oligomers can be used to modulate morphogenic processes such as
CC metastatic cancer as well as angiogenesis-dependent processes such as
CC reproduction, development, wound healing, tissue repair, and
CC angiogenesis-dependent diseases. Also, antisera directed to the Fab
regions of endostatin oligomer antibodies can be administered to block
the ability of endogenous endostatin oligomer antisera to bind endostatin
oligomers.

XX Sequence 182 AA;

Query Match	100.0%	Score 893	DB 21	Length 182
Best Local Similarity	100.0%	Pred. No. 3.6e-101		
Matches 170	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	VALNPLPSGGMGIGADPQCFQOARAVGLACTFPRAFLLSSRLQDLYSTRADRAAVPIV	60	
Db	13	VALNPLPSGGMGIGADPQCFQOARAVGLACTFPRAFLLSSRLQDLYSTRADRAAVPIV	72	
QY	61	NLKDELPPSWALPESGSEGPLKPGARIPSPGKDVLRHPPTWPKSVHGPSNRRRLTE	120	
Db	73	NLKDELPPSWALPESGSEGPLKPGARIPSPGKDVLRHPPTWPKSVHGPSNRRRLTE	132	
QY	121	SYCEWTRTEAPSPATGQASSLLGGRIILGSGAASCHNAVIVLCIENSPMTAS	170	
Db	133	SYCEWTRTEAPSPATGQASSLLGGRIILGSGAASCHNAVIVLCIENSPMTAS	182	
RESULT 7				
ID	AAU00897	AAU00897 standard; Protein; 182 AA.		
AC	AAU00897			
XX	04-JUL-2001	(first entry)		
DT	XX	Human Endostatin(TM) C-terminus minus 1 protein.		
DE	XX			
KW	XX	Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;		
KW	XX	blood borne tumour; leukaemia; tumour metastasis; benign tumour;		
KW	XX	haemangioma; acoustic neuroma; neurofibroma; trachoma; rubecosis;		
KW	XX	pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;		
KW	XX	ocular angiogenic disease; diabetic retinopathy; macular degeneration;		
KW	XX	retinopathy of prematurity; macular corneal graft rejection;		
KW	XX	neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;		
KW	XX	myocardial angiogenesis; plaque neovascularisation; telangiectasia;		
KW	XX	haemophilic joint; angiodioma; wound granulation; variant;		
KW	XX	C-terminus minus 1 protein.		
OS	XX	Homo sapiens.		
XX	XX	WO200119989-A2.		
PN	XX	22-MAR-2001.		
PD	XX			
PP	XX	14-SEP-2000; 2000MO-US25166.		
PR	XX	14-SEP-1999; 99US-0153698.		
XX	XX	(ENTR-) ENTREMED INC.		
PA	XX			
PI	XX	Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ,		
PI	XX	Bermejo JL, Mistry FR, Shepard SR, Schriber JL,		
DR	XX	WPI; 2001-244802/25.		
XX	XX	N-PSDB; AAS00897.		
PT	XX	Producing Endostatin protein for treating angiogenesis mediated		
PT	XX	diseases such as solid tumours, comprises recombinantly producing the		
PT	XX	protein using an expression system, and recovering and purifying the		
PT	XX	protein		
PS	XX	Claim 5, Page 30; 67pp; English.		
XX	XX			
CC	XX	The sequence represents Human Endostatin(TM) C-terminus minus 1		
CC	XX	protein, a natural variant lacking the C-terminal amino acid of		
CC	XX	Endostatin(TM) recovered from fermentations of <i>Pichia</i>		
CC	XX	<i>pastoris</i> cultures harbouring a expression plasmid containing the		
CC	XX	Endostatin(TM) DNA sequence given in AAS00867. The new method of the		
CC	XX	invention is useful for producing, recovering and purifying Endostatin		
CC	XX	(TM) from biological sources, such as biological fluids, tissues, cells,		
CC	XX	culture media, and fermentation media. Endostatin(TM) is useful for		
CC	XX	treating angiogenesis mediated diseases such as solid tumours, blood		
CC	XX	borne tumours, leukemias, tumour metastases, benign tumours, e.g.		

CC	haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
CC	granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
CC	e.g., diabetic retinopathy, retinopathy of prematurity, macular
CC	degeneration, corneal graft rejection, neovascular glaucoma, colon
CC	cancer, retrolental fibroplasia, rubeosis, Osler-Weber Syndrome,
CC	myocardial angiogenesis, plaque neovascularisation, telangiectasia,
CC	haemophillic joints, angiodioma, and wound granulation. Endostatin(TM
CC	is also useful for treating disease of excessive or abnormal stimulation
CC	of endothelial cells such as intestinal adhesions, atherosclerosis,
CC	scleroderma and hypertrophic scars. Higher yields of more purified, and
CC	biologically active Endostatin(TM) are obtained by the new method.
CC	Endostatin(TM) can be stored in buffers for extended periods of time, and
CC	also subjected to lyophilisation, while preserving biological activity.
CC	Centrifugation of broth from fermentation steps in production is avoided,
CC	preventing unwanted potential cellular lysis and contamination with
CC	additional proteins, pigments, enzymes and other cellular chemicals and
CC	debris.
CC	
XX	
XX	Sequence 182 AA;
XX	
XX	
XX	Query Match 100.0%; Score 893; DB 22; Length 182;
XX	Best Local Similarity 100.0%; Pred. No.3.6e-101; Indels 0; Gaps 0;
XX	Matches 170; Conservative 0; Mismatches 0;
QY	1 VALNSPLSGKRGIRGADPFCFOGARAVGLGTRAPFLSRLDLYSIVRADDAVAPIV 60
DB	13 VALNSPLSGKRGIRGADPFCFOGARAVGLGTRAPFLSRLDLYSIVRADDAVAPIV 72
QY	61 NLKDELFLPSSWEALFSGSEGPLKPGARIFSPDGKDVLRHPWPQKSVWGGSDPNGRRLTE 120
DB	73 NLKDELFLPSSWEALFSGSEGPLKPGARIFSPDGKDVLRHPWPQKSVWGGSDPNGRRLTE 132
QY	121 SYCEWTFREAPSPATGOASSLLGGRLLGQSASCHAAIVLVLCIENSFMPTAS 170
DB	133 SYCEWTFREAPSPATGOASSLLGGRLLGQSASCHAAIVLVLCIENSFMPTAS 182
XX	
XX	RESULT 8
XX	AAU77951
XX	ID AAU77951 standard; Protein: 182 AA.
XX	AC AAU77951;
XX	DT 02-JUL-2002 (first entry)
XX	DE Amino acid sequence for human endostatin.
XX	Human; immunoglobulin; anti-vascular endothelial growth factor antibody;
XX	anti-VEGF antibody; monoclonal antibody 2C3 ATCC PTA 1955; VEGF receptor;
XX	VEGFR2; KDR/PLK-1; VEGFR1; Flt-1; angiogenesis; macular degeneration;
XX	ocular neovascular disease; cancer; vascularised solid tumour; AIDS;
XX	metastatic tumour; endothelial cell proliferation; inflammatory disorder;
XX	atherosclerosis; diabetic retinopathy; corneal graft rejection;
XX	acquired immune deficiency syndrome; infection; restenosis; fungal ulcer;
XX	sickle cell anaemia; endometriosis; endostatin.
XX	
XX	Homo sapiens.
XX	
XX	PN AU200179401-A.
XX	
XX	PD 06-DEC-2001.
XX	
XX	PF 12-OCT-2001; 2001AU-0079401.
XX	
XX	PR 28-APR-2000; 2000AU-0048049.
XX	PR 12-OCT-2001; 2001AU-0079401.
XX	
XX	PA (TEXA) UNIV TEXAS SYSTEM.
XX	
XX	PI Thorpe PE, Brekken RA,
XX	
XX	DR WPI; 2002-281368/33.

PT Immunocjugate compositions for treating cancer by inhibiting
 PT angiogenesis and for delivering a diagnostic agent to tumour, comprises
 PT anti-vascular endothelial growth factor antibody attached to a
 PT biological agent -

PS Example 10; Page 12-13 (Sequence listing); 300pp; English.

XX
 CC The present invention relates to antibody-based compositions comprising
 CC an immunocjugate such as anti-vascular endothelial growth factor
 CC (VEGF) antibody (Ab) (or its antigen-binding fragment), attached to a
 CC biological agent, where the Ab binds to the same epitope as the
 CC monoclonal antibody (mAb) 2C3 ATCC PTA 1593, and significantly inhibits
 CC VEGF binding to the VEGF receptor VEGFR2 (KDR/Flk-1) without inhibiting
 CC VEGF binding to the VEGF receptor VEGFR1 (Flt-1). The compositions
 CC of the invention are useful in therapy, and diagnosis, for inhibiting
 CC angiogenesis in an animal having ocular neovascular disease or macular
 CC degeneration, and for delivering a biological agent to a vascularised
 CC tumour. The compositions can also be used for treating cancer and
 CC subjects at risk of developing, a vascularised solid tumour, a metastatic
 CC tumour or metastases from a primary tumour. The composition is useful
 CC for specifically inhibiting VEGF-induced endothelial cell proliferation,
 CC without significantly inhibiting VEGF-induced macrophage, osteoclast or
 CC chondrocyte function. The compositions can be used for treating various
 CC diseases such as inflammatory disorders, atherosclerosis, diabetic
 CC retinopathy, restenosis, acquired immune deficiency syndrome (AIDS),
 CC blood borne tumours, corneal graft rejection, Crohn's disease, fungal
 CC ulcers, infections, sickle cell anaemia, and endometriosis. The present
 CC sequence represents human endostatin. Endostatin may be attached or
 CC functionally associated with anti-VEGF antibodies.

XX Sequence 182 AA;

Query Match 100.0%; Score 893; DB 23; Length 182;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
 DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWELFSGSEGPLKPGARIFSPDGKDVLRHPTWPKQSVWHSDENGRRLTE 120
 DB 73 NLKDELLFPSEWELFSGSEGPLKPGARIFSPDGKDVLRHPTWPKQSVWHSDENGRRLTE 132

QY 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
 DB 133 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 9

AAV08693
 ID AAV08693 standard; Protein; 183 AA.

XX AAV08693;

XX 10-AUG-1999 (first entry)

XX Human endostatin protein fragment.

XX KW Plasminogen; human; angiotensin; endostatin; gene therapy; vector;
 KW anti-angiogenic; attenuation; cytosolic; anti-diabetic; ophthalmology;
 KW tumour growth; solid tumour; diabetic retinopathy; retina.

XX Homo sapiens.

XX WO9926480-A1.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-US24950.

XX 20-NOV-1997; 97US-0975424.

PA (GENE-) GENETIX PHARM INC.
 PA (MAST) MASSACHUSETTS INST TECHNOLOGY.

XX Bachelot T, Leboulch P, Pawluk RJ;

XX WPI, 1999-357696/30.

XX N-PSDB; AAX77719.

PT Anti-angiogenic gene therapy vectors

PS Disclosure; Page 74-75; 83pp; English.

XX
 CC This invention describes a novel viral gene therapy vector comprising a
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
 CC from human or murine angiotensin, human or murine endostatin and
 CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
 CC sufficiently attenuated for use in human gene therapy. The products of
 CC the invention have anti-angiogenic, cytosolic, anti-diabetic and
 CC ophthalmological activity. The vector is used in gene therapy for
 CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
 CC expresses an anti-angiogenic polypeptide. An additional use comprises
 CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
 CC inhibits angiogenesis in the vicinity of the retina. The vector is
 CC administered to cells ex vivo and then administered to the patient.

XX Sequence 183 AA;

Query Match 100.0%; Score 893; DB 20; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
 DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWELFSGSEGPLKPGARIFSPDGKDVLRHPTWPKQSVWHSDENGRRLTE 120
 DB 73 NLKDELLFPSEWELFSGSEGPLKPGARIFSPDGKDVLRHPTWPKQSVWHSDENGRRLTE 132

QY 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
 DB 133 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 10

AAV02113
 ID AAV02113 standard; Protein; 183 AA.

XX AAV02113;

XX 16-JUL-1999 (first entry)

XX SEQ ID 76 of WO9916889.

XX Angiotensin; endostatin; interferon; thrombospondin;
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;
 KW anti-tumour; multifunctional protein; angiogenic-mediated disease;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW tumor cell production.

XX Homo sapiens.

XX WO9916889-A1.

XX 08-APR-1999.

XX 30-SEP-1998; 98WO-US20464.

XX 01-OCT-1997; 97US-0060609.

XX (SEAR) SEARLE & CO G D.

XX Bolarowski MA, Caparon MH, Casperon GF, Gregory SA;

XX PI

PI Klein BK, McKeam JP;
XX MPI, 1999-255098/21.
DR
XX
PT New multifunctional proteins useful for treating angiogenic-mediated
PT diseases
XX
PS Disclosure, Page 106-107, 121pp; English.
XX
CC The specification describes multifunctional proteins which comprise
CC combinations of angiostatin, endostatin, interferon, thrombospondin,
CC interferon-inducible protein and platelet factor 4, and have
CC anti-angiogenic and/or anti-tumor activity. The multifunctional protein
CC may exhibit useful properties such as having similar or greater
CC biological activity when compared to a single factor or by having
CC improved half-life or decreased adverse side effects, or a combination
CC of these properties. The proteins can be used for treating an
CC angiogenic-mediated disease, e.g. cancer, diabetic retinopathy, macular
CC degeneration, or arthritis. They can also be used for inhibiting the
CC production of tumor cells (characteristic of lung, breast, ovarian,
CC prostate, pancreatic, gastric, colon, renal, bladder cancers; melanoma,
CC hepatoma, sarcoma and lymphoma) in a patient and for inhibiting tumor
CC growth. The present sequence is used in the course of the invention.
XX
SQ Sequence 183 AA;
Query Match 100.0%; Score 893; DB 20; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.6e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAVPIV 72
QY 61 NLKDELFPSEWALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWHSDDPNGRLTE 120
DB 73 NLKDELFPSEWALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWHSDDPNGRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 182
RESULT 11
AAB30493
ID AAB30493 standard; Protein; 183 AA.
XX
AC AAB30493;
XX
DT 06-MAR-2001 (first entry)
XX
DE Amino acid sequence of human endostatin encoded by plasmid pMALch#15.
XX
KM Streptomyces sp. strain C5; SmuA; S. venezuelae; alpha-amylase;
KW endostatin; cancer; tumour growth; angiogenesis.
XX
OS Homo sapiens.
XX
PN WO200060945-A1.
XX
PD 19-OCT-2000.
XX
PF 12-APR-2000; 2000WO-US09747.
XX
PR 13-APR-1999; 99US-0129084.
XX
PA (MERI) MERCK & CO INC.
XX
PI Desanti CL, Strohl WR;
XX
DR MPI; 2000-686970/67.
XX
N-PSDB; AAC62023.
XX

PT Preparation of soluble recombinant endostatin involves transforming
PT Streptomyces host with expression vector comprising nucleotide
PT sequence encoding endostatin operably linked to linker and leader
PT peptide
XX
PS Example 1; Fig 6; 57pp; English.
XX
CC The present sequence represents human endostatin. The protein is
CC expressed in Streptomyces. Leader sequences of Streptomyces sp. strain
CC C5 SmuA and S. venezuelae alpha-amylase proteins are linked to the
CC C-terminal of endostatin. This ensures that endostatin protein is
CC produced as a secreted, soluble protein which needs no refolding, is
CC stable in the fermentation broth and is produced in large quantities.
CC The method is used for preparing soluble recombinant human, murine or
CC primate endostatin, which is useful in the treatment of cancer,
CC inhibition of tumour growth, inhibition of angiogenesis, isolation of
CC receptors for endostatin and for identification of anti-angiogenic
CC compounds in assays. The endostatin protein is produced as a secreted,
CC soluble protein which needs no refolding, is stable in the fermentation
CC broth and is produced in large quantities. Streptomyces are amenable
CC for cultivation in large fermentations allowing for large quantities of
CC soluble endostatin to be produced.
XX
SQ Sequence 183 AA;
Query Match 100.0%; Score 893; DB 21; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.6e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAVPIV 72
QY 61 NLKDELFPSEWALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWHSDDPNGRLTE 120
DB 73 NLKDELFPSEWALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWHSDDPNGRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 182
RESULT 12
AAB16451
ID AAB16451 standard; Protein; 183 AA.
XX
AC AAB16451;
XX
DT 27-OCT-2000 (first entry)
XX
DE Human endostatin protein sequence.
XX
KM Angiogenesis-inhibiting protein receptor; angiogenesis; angiostatin;
KW endostatin; plasmogen; laminin; treatment; wound healing; solid tumour;
KW psoriasis; scleroderma; myocardial angiogenesis; Crohn's disease;
KW cerebral collateral; arteriovenous malformation; rhepsis; cancer;
KW diabetic retinopathy; arthritis; wound healing; peptic ulcer;
KW Helicobacter related disease; fracture; cat scratch fever.
XX
OS Homo sapiens.
XX
PN WO200032631-A2.
XX
PD 08-JUN-2000.
XX
PF 06-DEC-1999; 99WO-US28897.
XX
PR 04-DEC-1998; 98US-0206059.
XX
PA (ENTR-) ENTREMED INC.
XX
PI Macdonald NJ, Sim KL;
XX

DR WPI; 2000-412290/35.
 XX
 PT New angiogenesis-inhibiting protein receptors, useful in methods for
 PT treating diseases and processes that are mediated by angiogenesis, such
 PT as solid tumours, psoriasis, scleroderma and myocardial angiogenesis -
 XX
 XX Disclosure; Figure 3; 100pp; English.
 PS
 CC This invention relates to angiogenesis-inhibiting protein receptors, and
 CC the DNA sequences encoding them. Angiogenesis is the generation of new
 CC blood vessels into a tissue, and normally occurs in wound healing,
 CC foetal and embryonal development and the formation of the corpus luteum,
 CC endometrium and placenta. Angiostatin is a protein (see AAB16450 and
 CC AAB68202) involved in angiogenesis, and has an amino acid sequence
 CC similar to that of a plasminogen fragment (see murine plasminogen
 CC AAB16490). Angiostatin has the ability to inhibit angiogenesis.
 CC Endostatin is also an angiogenesis inhibiting protein (see AAB16451 and
 CC AAB68203). Sequences AAB68242 and AAB16522 represent coding and protein
 CC sequences of human laminin. Laminin is an angiostatin binding protein,
 CC and some of the peptides of the invention share homology with regions of
 CC laminin. Peptides AAB16452-B16521 (excluding AAB16490) are the
 CC angiogenesis-inhibiting protein receptor fragments of the invention. The
 CC peptides bind either angiostatin or endostatin and can be used in methods
 CC for treating diseases and processes that are mediated by angiogenesis,
 CC such as solid tumours, psoriasis, scleroderma, myocardial angiogenesis,
 CC Crohn's disease, cerebral collaterals, arteriovenous malformations,
 CC rheiosis, diabetic retinopathy, arthritis, wound healing, peptic ulcers,
 CC Helicobacter related diseases, fractures, placentaion and cat scratch
 CC fever. They are useful for the detection and prognosis of cancer. DNA
 CC sequences A628204-A628241 encode the peptides of the invention.
 CC
 XX
 SQ Sequence 183 AA;
 Query Match 100.0%; Score 893; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101; Indels 0; Gaps 0;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSRLDDLSIVRRADRAAVPIV 60
 DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSRLDDLSIVRRADRAAVPIV 72
 QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHSGSDPNGRLTE 120
 DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHSGSDPNGRLTE 132
 QY 121 SYCETWTEAPSATGQASSILGRLGQSAASCHAAIYIVLCIENSFMTAS 170
 DB 133 SYCETWTEAPSATGQASSILGRLGQSAASCHAAIYIVLCIENSFMTAS 182

RESULT 13
 AAY90771
 ID AAY90771 standard; Protein; 183 AA.
 XX
 AC AAY90771;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human angiogenesis inhibiting factor 1 protein.
 XX
 KW Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;
 KW abnormal vessel disease.
 XX
 OS Homo sapiens.
 XX
 PN CN1244536-A.
 XX
 PD 16-FEB-2000.
 XX
 PF 10-AUG-1998; 98CN-0117150.
 XX
 PR 10-AUG-1998; 98CN-0117150.
 XX

PA (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.
 XX
 PI Yang Z, Guo W;
 XX
 DR WPI; 2000-38168/34.
 DR N-PsDB; AAB29884.
 XX
 XX Angiogenesis inhibiting factor 1 and its derivative useful for treating
 PT tumors -
 PT
 PS Claim 1; Fig 5; 41pp; Chinese.
 XX
 CC The present sequence represents an angiogenesis inhibiting factor (1),
 CC designated IAF-1. The present invention also describes: (1) preparation
 CC of (1) and its derivative; (2) an IAF binding acceptor and its
 CC preparation; and (3) an IAF antibody. (1) is useful for preparing new
 CC biological preparations for effectively treating various tumors and
 CC abnormal-vessel diseases. The IAF antibody is preferably a polyclonal
 CC antibody, mosaic antibody, single stranded antibody and human originated
 CC antibody.
 CC
 XX
 SQ Sequence 183 AA;
 Query Match 100.0%; Score 893; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101; Indels 0; Gaps 0;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSRLDDLSIVRRADRAAVPIV 60
 DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSRLDDLSIVRRADRAAVPIV 72
 QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHSGSDPNGRLTE 120
 DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHSGSDPNGRLTE 132
 QY 121 SYCETWTEAPSATGQASSILGRLGQSAASCHAAIYIVLCIENSFMTAS 170
 DB 133 SYCETWTEAPSATGQASSILGRLGQSAASCHAAIYIVLCIENSFMTAS 182

RESULT 14
 AAY70252
 ID AAY70252 standard; Protein; 183 AA.
 XX
 AC AAY70252;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Human angiogenesis inhibitor, endostatin.
 XX
 KW Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin;
 KW angiogenesis; inhibitor; cytosolic; antirheumatoid; antiarteritic;
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;
 KW vasotropic; vulnerary; treatment; antiarteriosclerosis; tumour;
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW wound granulation; keloid scar; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200011033-A2.
 XX
 PD 02-MAR-2000.
 XX
 PF 25-AUG-1999; 99WO-US19329.
 XX
 PR 25-AUG-1998; 98US-0097883.
 XX
 PA (LEXI-) LEXINGEN PHARM CORP.
 XX
 PI Lo K, Li Y, Gillies SD;
 XX

DR WPI: 2000-237616/20.
 DR N-PSDB; AAS51291.
 PT Novel fusion protein of angiotensin or endostatin and an immunoglobulin
 PC region, useful for treating conditions mediated by angiogenesis,
 PT such as rheumatoid arthritis, tumors and macular degeneration -
 XX
 PS Example 1; Pages 41-42; 68pp; English.
 CC The patent discloses a DNA molecule encoding a fusion protein comprising
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 CC inhibitor selected from angiotensin, endostatin, a plasminogen fragment
 CC having angiotensin activity, a collagen XVII fragment having endostatin
 CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumors,
 CC blood born tumors, tumor metastasis, benign tumors including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasis, haemophilic joints,
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a human endostatin used in the
 CC construction of immunofusin containing human immunoglobulin gamma
 CC (Igg) Fc fragment.
 CC
 CC Sequence 183 AA;
 SQ
 Query Match 100.0%; Score 893; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLVSIYRRADRAVPIV 60
 DB 13 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLVSIYRRADRAVPIV 72
 QY 61 NLKDELLPFSWEALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKQSVHGSDPNGRRLTE 120
 DB 73 NLKDELLPFSWEALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKQSVHGSDPNGRRLTE 132
 QY 121 SYCETRTAPSPATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
 DB 133 SYCETRTAPSPATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182
 RESULT 15
 AAU00896
 ID AAU00896 standard; Protein; 183 AA.
 AC AAU00896;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human Endostatin(TM) protein.
 XX
 KW Human, Endostatin(TM); angiogenesis mediated disease; solid tumors;
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
 KW retinopathy of prematurity; macular corneal graft rejection;
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
 KW haemophilic joint; angiofibroma; wound granulation.
 XX
 OS Homo sapiens.
 XX
 XX WO200119989-A2.
 XX

PD 22-MAR-2001.
 XX
 XX 14-SEP-2000; 2000WO-US25166.
 XX
 PR 14-SEP-1999; 99US-0153698.
 XX
 XX (ENTR-) EXTREMED INC.
 XX
 XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madgen J, Boerner RJ;
 PI Bernejo LL, Mistry FR, Shepard SR, Schrimsher JL;
 DR WPI: 2001-244802/25.
 DR N-PSDB; AAS00867.
 CC Producing Endostatin protein for treating angiogenesis mediated
 CC diseases such as solid tumors, comprises recombinantly producing the
 CC protein using an expression system, and recovering and purifying the
 CC protein -
 CC
 CC Claim 5; Page 29; 67pp; English.
 CC The sequence represents Human Endostatin(TM). The new method of the
 CC invention is useful for producing, recovering and purifying Endostatin
 CC (TM) from biological sources, such as biological fluids, tissues, cells,
 CC culture media, and fermentation media. Endostatin(TM) is useful for
 CC treating angiogenesis mediated diseases such as solid tumours, blood
 CC borne tumors, leukemias, tumour metastases, benign tumors, e.g.
 CC haemangioma, acoustic neuromas, neurofibromas, trachomas, and pyogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases,
 CC e.g., diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, colon
 CC cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome,
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,
 CC haemophilic joints, angiofibroma, and wound granulation. Endostatin(TM)
 CC is also useful for treating disease of excessive or abnormal stimulation
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and
 CC biologically active Endostatin(TM) are obtained by the new method.
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and
 CC also subjected to lyophilisation, while preserving biological activity.
 CC Centrifugation of broth from fermentation steps in production is avoided,
 CC preventing unwanted potential cellular lysis and contamination with
 CC additional proteins, pigments, enzymes and other cellular chemicals and
 CC debris.
 CC
 CC Sequence 183 AA;
 SQ
 Query Match 100.0%; Score 893; DB 22; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3.6e-101;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLVSIYRRADRAVPIV 60
 DB 13 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLVSIYRRADRAVPIV 72
 QY 61 NLKDELLPFSWEALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKQSVHGSDPNGRRLTE 120
 DB 73 NLKDELLPFSWEALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKQSVHGSDPNGRRLTE 132
 QY 121 SYCETRTAPSPATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
 DB 133 SYCETRTAPSPATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182
 Search completed: February 11, 2003, 20:41:26
 Job time : 38 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:37:18 ; Search time 12 Seconds

(without alignments)
587,581 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893
Sequence: 1 VALNSPLSGMRGIRGADFO.....ASCHNAVYLCTENSEPTAS 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	1516	1 CAIH_HUMAN	P39060 homo sapien
2	778	87.1	1527	1 CAIH_MOUSE	P39061 mus musculu
3	504	56.4	1388	1 CAIE_HUMAN	P39059 homo sapien
4	76	8.5	556	1 GLI_CHICK	P55878 gallus gall
5	75	8.4	271	1 Y4BG_RHISN	P55374 rhizobium s
6	74	8.3	296	1 CRK_XENLA	P87378 xenopus lae
7	74	8.3	780	1 STRN_RAT	P70483 rattus norv
8	73.5	8.2	1653	1 YFHH_ECOLI	P76578 escherichia
9	73	8.2	512	1 YP67_ECOLI	P29014 escherichia
10	72.5	8.1	884	1 YP67_MYTCTU	P50654 mycobacteri
11	72	8.1	304	1 CRK_HUMAN	P46108 homo sapien
12	72	8.1	304	1 CRK_MOUSE	P64010 mus musculu
13	72	8.1	304	1 CRK_RAT	P63768 rattus norv
14	72	8.1	613	1 HS75_CANAL	P87222 candida alb
15	70	7.8	780	1 STRN_MOUSE	P55106 mus musculu
16	69	7.7	266	1 CB21_SINAI	P13851 sinapis alb
17	69	7.7	309	1 YHCC_ECOLI	P45476 escherichia
18	69	7.7	326	1 TMOE_PSEME	P00460 pseudomonas
19	69	7.7	462	1 MYCN_MOUSE	P03966 mus musculu
20	69	7.7	1233	1 NMBS_HUMAN	P14957 homo sapien
21	68.5	7.7	390	1 PGK_BUCAI	P57525 buchera ap
22	68.5	7.7	1289	1 CSAB_BACUD	P45753 bacillus th
23	68	7.6	953	1 CSAB_HUMAN	P09239 homo sapien
24	68	7.6	999	1 MERK_HUMAN	P12866 homo sapien
25	67.5	7.6	348	1 NUTN_BRARE	P04159 brachydanio
26	67.5	7.6	536	1 CAR2_HUMAN	P09257 homo sapien
27	67.5	7.6	579	1 PAAR_MOUSE	P008914 mus musculu
28	67.5	7.6	692	1 GYRB_BARBA	P94281 bartonella
29	67	7.5	145	1 ANP_RANCA	P18909 rana catesb
30	67	7.5	250	1 HXB5_HUMAN	P17482 homo sapien
31	67	7.5	354	1 YM28_MYTCTU	P10512 mycobacteri
32	67	7.5	577	1 BAG3_MOUSE	P09111 mus musculu
33	66.5	7.4	368	1 GALT_HUMAN	P06755 homo sapien

34	66.5	7.4	449	1 DHE2_CLOST	P24295 clostridium
35	66.5	7.4	1548	1 SMCY_MOUSE	P06240 mus musculu
36	66	7.4	250	1 HXB9_MOUSE	P20615 mus musculu
37	66	7.4	342	1 G3P1_ANAVA	P34916 anabaena va
38	66	7.4	399	1 SUCC_CAUCR	P04894 caulobacter
39	66	7.4	419	1 P4TK_PSECL	P31521 pseudomonas
40	66	7.4	484	1 TRPG_YEAST	P00937 saccharomyc
41	66	7.4	505	1 CH01_HUMAN	P09236 homo sapien
42	66	7.4	622	1 PPCC_MOUSE	P09234 mus musculu
43	66	7.4	713	1 ACS1_YEAST	P01574 saccharomyc
44	66	7.4	876	1 AREA_YEAST	P17429 emeticocella
45	66	7.4	3491	1 ERY1_SACER	P03131 saccharopol

ALIGNMENTS

RESULT 1	ID	CAIH_HUMAN	STANDARD	PRT	1516 AA.
AC	P39060	Q9Y608; Q9Y607; Q9UK38;			
DT	01-FEB-1995	(Rel. 31. Created)			
DT	15-OCT-2001	(Rel. 40. Last sequence update)			
DT	15-JUN-2002	(Rel. 41. Last annotation update)			
DE	Collagen alpha 1(XVII) chain precursor [Contains: Endostatin].				
GN	COL18A1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=98164096; PubMed=9503365;				
RX	Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;				
RA	"Complete primary structure of two variant forms of human type XVIII				
RT	collagen and tissue-specific differences in the expression of the				
RT	corresponding transcripts.";				
RL	Matrix Biol. 16:319-328(1998).				
RN	[2]	SEQUENCE FROM N.A.			
RP	MEDLINE=20289799; PubMed=10830953;				
RX	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,				
RA	Ohki H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,				
RA	Parik M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,				
RA	Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,				
RA	Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,				
RA	Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,				
RA	Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,				
RA	Minoshima S., Shimizu N., Nordle G., Hornischer K., Brandt P.,				
RA	Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,				
RA	Ramer J., Beck A., Klages S., Hennig S., Riesemann L., Dagand E.,				
RA	Wehrmeyer S., Borzym K., Gardiner K., Niretic D., Francis P.,				
RA	Lehrach H., Reinhardt R., Yaspo M.-L.;				
RT	"The DNA sequence of human chromosome 21.";				
RL	Nature 405:311-319(2000).				
RN	[3]	SEQUENCE OF 834-1516 FROM N.A.			
RP	MEDLINE=94245237; PubMed=8188291;				
RX	Oh S.-P., Warman M.L., Seidlin M.F., Cheng S., Knoll J.H., Timmons S.,				
RA	Olsen B.R.;				
RT	"Cloning of cDNA and genomic DNA encoding human type XVIII collagen				
RT	and localization of the alpha 1(XVIII) collagen gene to mouse				
RT	chromosome 10 and human chromosome 21.";				
RL	Genomics 19:494-499(1994).				
RN	[4]	SEQUENCE OF 1334-1516 FROM N.A.			
RP	TISSUE=Placenta;				
RC	Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.;				
RA	"Cloning and expression of human endostatin gene in Escherichia				
RT	coli.";				
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]	INVOLVEMENT IN KNOBLOCH SYNDROME.			

RX MEDLINE=20400145; PubMed=10942434;
 RA Sertie A.L., Soares F., Camargo A.A., Zatz M., Brahe C.,
 RA Passos-Bueno M.R.;
 RT "Collagen XVII, containing an endogenous inhibitor of angiogenesis
 RT and tumor growth, plays a critical role in the maintenance of retinal
 RT structure and in neural tube closure.";
 RL Hum. Mol. Genet. 9:2051-2058(2000).
 RN [6]
 RP VARIANT ASN-1437.
 RX MEDLINE=21518361; PubMed=11606364;
 RA Inghetti P., Suzuki O., Godel P.H., Alves V.A., Sertie A.L.,
 RA Zorick T., Soares F., Camargo A., Moreira E.S., di Loreto C.,
 RA Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.;
 RT "A polymorphism in endostatin, an angiogenesis inhibitor, predisposes
 RT for the development of prostatic adenocarcinoma.";
 RL Cancer Res. 61:7375-7378(2001).
 CC -1- FUNCTION: COL1A8 PROBABLY PLAYS A MAJOR ROLE IN DETERMINING THE
 CC RETINAL STRUCTURE AS WELL AS IN THE CLOSURE OF THE NEURAL TUBE.
 CC -1- FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
 CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
 CC BINDING TO THE HEPARAN SULPHATE PROTEOGLYCANS INVOLVED IN GROWTH
 CC FACTOR SIGNALING.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A SHORT FORM/NC1-303 AND A LONG
 CC FORM/NC-493 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PRESENT IN MULTIPLE ORGANS WITH HIGHEST LEVELS
 CC IN LIVER, LUNG AND KIDNEY.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- POLYMORPHISM: There is an association between a polymorphism in
 CC position 1437 and prostate cancer. Heterozygous Asn-1437
 CC individuals have a 2.5 times increased chance of developing
 CC prostate cancer as compared with homozygous Asp-1437 individuals.
 CC -1- DISEASE: Defects in COL1A1 are a cause of Knobloch syndrome (KS
 CC or KNO); an autosomal recessive disorder defined by the occurrence
 CC of high myopia, vitreoretinal degeneration with retinal
 CC detachment, macular abnormalities and occipital encephalocele.
 CC -1- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
 CC INTERRUPTED HELICES (FACIT) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC
 DR EMBL; AF018081; AAC39658.1; -;
 DR EMBL; AF018082; AAC39659.1; -;
 DR EMBL; AL163302; CAB90482.1; -;
 DR EMBL; L22548; AAA51864.1; -;
 DR EMBL; AF184060; AAF01310.1; ALT_INIT.
 DR HSSP; P39061; IKOE.
 DR GlycoSiteDB; P39060; -;
 DR Genew; HGNC:2195; COL18A1.
 DR MIM; 120328; -;
 DR MIM; 267750; -;
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR003129; TSEN.
 DR Pfam; PF01391; Collagen; 7.
 DR Pfam; PF02210; TSPN; 1.
 DR SMART; SM00282; Lamc; 1.
 DR SMART; SM00210; TSPN; 1.
 DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 DR Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
 DR Polymorphism.
 KW SIGNAL.
 FT CHAIN 1 23 POTENTIAL.
 FT CHAIN 24 1516 COLLAGEN ALPHA 1(XVII) CHAIN.
 FT CHAIN 1334 1516 ENDOSTATIN.
 FT CHAIN 24 516 NONHELICAL REGION 1 (NC1).
 FT CHAIN 517 550 TRIPLE-HELICAL REGION 1 (COL1).
 FT DOMAIN 551 560 NONHELICAL REGION 2 (NC2).

FT	DOMAIN	561	640	TRIPLE-HELICAL REGION 2 (COL2).
FT	DOMAIN	641	664	NONHELICAL REGION 3 (NC3).
FT	DOMAIN	665	786	TRIPLE-HELICAL REGION 3 (COL3).
FT	DOMAIN	787	809	NONHELICAL REGION 4 (NC4).
FT	DOMAIN	810	892	TRIPLE-HELICAL REGION 4 (COL4).
FT	DOMAIN	893	906	NONHELICAL REGION 5 (NC5).
FT	DOMAIN	907	948	TRIPLE-HELICAL REGION 5 (COL5).
FT	DOMAIN	949	961	NONHELICAL REGION 6 (NC6).
FT	DOMAIN	962	1034	TRIPLE-HELICAL REGION 6 (COL6).
FT	DOMAIN	1035	1044	NONHELICAL REGION 7 (NC7).
FT	DOMAIN	1045	1077	TRIPLE-HELICAL REGION 7 (COL7).
FT	DOMAIN	1078	1089	NONHELICAL REGION 8 (NC8).
FT	DOMAIN	1090	1111	TRIPLE-HELICAL REGION 8 (COL8).
FT	DOMAIN	1112	1118	NONHELICAL REGION 9 (NC9).
FT	DOMAIN	1119	1173	TRIPLE-HELICAL REGION 9 (COL9).
FT	DOMAIN	1174	1186	NONHELICAL REGION 10 (NC10).
FT	DOMAIN	1187	1204	TRIPLE-HELICAL REGION 10 (COL10).
FT	DOMAIN	1205	1516	NONHELICAL REGION 11 (NC11).
FT	CARBOHYD	68	68	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	129	129	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	164	164	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	691	691	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1329	1329	O-LINKED (GALNAC. . .) (POTENTIAL).
FT	DISULFID	1366	1506	/FTD-CAR-000150.
FT	DISULFID	1468	1498	BY SIMILARITY.
FT	SITE	1095	1097	CELL ATTACHMENT SITE (POTENTIAL).
FT	VARSPPLIC	1	180	MISSING (IN SHORT ISOFORM).
FT	VARSPPLIC	181	215	HTTAGCTLPAPPSPPSLCPAPPAAPVPPSS -> MA PQFWPMPRRRLDVLAVLVLGVRAASAP (IN SHORT ISOFORM).
FT	VARIANT	1437	1437	D -> N.
FT	CONFLICT	428	428	/FTID=VAR 012709.
FT	CONFLICT	841	841	F -> S (IN REF. 2).
FT	CONFLICT	877	877	I -> V (IN REF. 2).
FT	CONFLICT	886	886	V -> L (IN REF. 3).
FT	CONFLICT	912	912	P -> R (IN REF. 3).
FT	CONFLICT	933	933	P -> L (IN REF. 3).
FT	CONFLICT	975	975	P -> L (IN REF. 3).
FT	CONFLICT	1064	1064	A -> P (IN REF. 3).
FT	CONFLICT	1084	1084	L -> K (IN REF. 3).
FT	CONFLICT	1120	1120	P -> A (IN REF. 3).
FT	CONFLICT	1123	1123	P -> A (IN REF. 3).
FT	CONFLICT	1126	1126	P -> PGP (IN REF. 2).
FT	CONFLICT	1206	1206	G -> GQ (IN REF. 3).
FT	CONFLICT	1304	1304	R -> G (IN REF. 3).
FT	CONFLICT	1314	1314	A -> G (IN REF. 3).
FT	CONFLICT	1324	1324	LR -> CG (IN REF. 3).
FT	CONFLICT	1443	1443	R -> T (IN REF. 4).
FT	CONFLICT	1483	1483	S -> Y (IN REF. 4).
SO	SEQUENCE	1516 AA;	153840 MW;	3C70F29A447EE66 CRC64;

Query Match 100.0%; Score 893; DB 1; Length 1516;
 Best Local Similarity 100.0%; Pred. No. 5.6e-81;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VALNSPLSGMGRGIRGADPCFOGARAVGLAGFRAPFLSRLDLSIYRARAAVPIV	60
DB	1346	VALNSPLSGMGRGIRGADPCFOGARAVGLAGFRAPFLSRLDLSIYRARAAVPIV	1405
QY	61	NLDELFLPFSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTPKSVHSGSDPGRRLTE	120
DB	1406	NLDELFLPFSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTPKSVHSGSDPGRRLTE	1465
QY	121	SYCEWTRETPASATGQASLLGRLIGQSAASHAYIVLCINSEMTAS	170
DB	1466	SYCEWTRETPASATGQASLLGRLIGQSAASHAYIVLCINSEMTAS	1515

RESULT 2
 CAH MOUSE
 ID CAH_MOUSE STANDARD; PRT; 1527 AA.

AC P39061; 062002; 061437;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(XVII) chain precursor [Contains: Endostatin].
GN COL18A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (SHORT ISOFORM).
RC STRAIN=BALB/c; TISSUE=Liver;
RA MEDLINE=94245707; PubMed=818673;
RA Reh M.V., Hintsikka E., Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of mouse type XVII collagen,
RT partial structure of the corresponding gene, and comparison of the
RT alpha 1(XVII) chain with its homologue, the alpha 1(XV) collagen
RT chain.";
RL J. Biol. Chem. 269:13929-13935(1994).
RN [2]
RN SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA Reh M., Hintsikka E., Pihlajaniemi T.;
RT "Characterization of the mouse gene for the alpha-1 chain of type
RT XVII collagen (COL18A1) reveals that the three variant N-terminal
RT polypeptide forms are transcribed from two widely separated
RT promoters.";
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 213-1140 FROM N.A. (SHORT ISOFORM).
RA MEDLINE=94240112; PubMed=8183894;
RA Reh M.V., Pihlajaniemi T.;
RT "Alpha 1(XVII), a collagen chain with frequent interruptions in the
RT collagenous sequence, a distinct tissue distribution, and homology
RT with type XV collagen.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).
RN [4]
RN SEQUENCE OF 240-1527 FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=94240111; PubMed=8183893;
RA Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Ooshima A., Olsen B.R.;
RT "Isolation and sequencing of cDNAs for proteins with multiple domains
RT of Gly-Xaa-Yaa repeats identify a distinct family of collagenous
RT proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233(1994).
RN [5]
RN CHARACTERIZATION OF ENDOSTATIN, AND PARTIAL SEQUENCE.
RA MEDLINE=97160848; PubMed=9008166;
RA O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,
RA Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;
RT "Endostatin: an endogenous inhibitor of angiogenesis and tumor
RT growth.";
RL Cell 88:277-285(1997).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.
RA MEDLINE=98169382; PubMed=9501087;
RA Hohenester E., Sasaki T., Olsen B.R., Timpl R.;
RT "Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A
RT resolution.";
RL EMBO J. 17:1656-1664(1998).
RN [7]
RN FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
RN PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
RN BINDING TO THE HEPARAN SULPHATE PROTEGLYCANS INVOLVED IN GROWTH
RN FACTOR SIGNALING.
RN [8]
RN ALTERNATIVE PRODUCTS. 2 ISOFORMS, A SHORT FORM AND A LONG FORM
RN (SHOWN HERE), ARE PRODUCED BY ALTERNATIVE SPLICING.
RN [9]
RN UNIT: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
RN UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
RN [10]
RN SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
RN INTERRUPTED HELICES (FACIT) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-eb.ch/announce/>
CC or send an email to license@isb-eb.ch).
CC -----
DR EMBL; L16898; AAA37434.1; -
DR EMBL; U03714; AAA20657.1; -
DR EMBL; U03715; AAC52901.1; -
DR EMBL; U34606; AAC52901.1; JOINED.
DR EMBL; U34608; AAC52901.1; JOINED.
DR EMBL; U34609; AAC52901.1; JOINED.
DR EMBL; U34610; AAC52901.1; JOINED.
DR EMBL; U34611; AAC52901.1; JOINED.
DR EMBL; U34612; AAC52901.1; JOINED.
DR EMBL; U34613; AAC52901.1; JOINED.
DR EMBL; U03716; AAC52902.1; JOINED.
DR EMBL; U03718; AAC52901.1; JOINED.
DR EMBL; U03715; AAC52902.1; -
DR EMBL; U34607; AAC52902.1; JOINED.
DR EMBL; U34608; AAC52902.1; JOINED.
DR EMBL; U34609; AAC52902.1; JOINED.
DR EMBL; U34610; AAC52902.1; JOINED.
DR EMBL; U34611; AAC52902.1; JOINED.
DR EMBL; U34612; AAC52902.1; JOINED.
DR EMBL; U03716; AAC52902.1; JOINED.
DR EMBL; U03718; AAC52902.1; JOINED.
DR EMBL; U11636; AAC52178.1; -
DR EMBL; L22545; AAA19787.1; -
DR PDB; 1XOE; 16-FEB-99.
DR MGI; MGI:88451; Coll18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen_8.
DR Pfam; PF02210; TSPN; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
KW 3D-structure.
FT CHAIN 1 26
FT SIGNAL 27 1527
FT CHAIN 1344 1527
FT DOMAIN 27 538
FT DOMAIN 539 565
FT DOMAIN 566 575
FT DOMAIN 576 649
FT DOMAIN 650 673
FT DOMAIN 674 795
FT DOMAIN 796 818
FT DOMAIN 819 901
FT DOMAIN 902 915
FT DOMAIN 916 957
FT DOMAIN 958 970
FT DOMAIN 971 1043
FT DOMAIN 1044 1053
FT DOMAIN 1054 1066
FT DOMAIN 1067 1098
FT DOMAIN 1099 1122
FT DOMAIN 1123 1129
FT DOMAIN 1130 1181
FT DOMAIN 1182 1194
FT DOMAIN 1195 1212
FT DOMAIN 1213 1527
FT CARBOHYD 338 338
FT CARBOHYD 700 700
FT DISULFID 1376 1516
FT DISULFID 1478 1508
FT SITE 1104 1106
FT SITE 1106 212
FT VARSPPLIC 1 238
FT VARSPPLIC 213 238
CC -----
CC POTENTIAL.
CC COLLAGEN ALPHA 1(XVII) CHAIN.
CC ENDOSTATIN.
CC NONHELICAL REGION 1 (NC1).
CC TRIPLE-HELICAL REGION 1 (COL1).
CC NONHELICAL REGION 2 (NC2).
CC TRIPLE-HELICAL REGION 2 (COL2).
CC NONHELICAL REGION 3 (NC3).
CC TRIPLE-HELICAL REGION 3 (COL3).
CC NONHELICAL REGION 4 (NC4).
CC TRIPLE-HELICAL REGION 4 (COL4).
CC NONHELICAL REGION 5 (NC5).
CC TRIPLE-HELICAL REGION 5 (COL5).
CC NONHELICAL REGION 6 (NC6).
CC TRIPLE-HELICAL REGION 6 (COL6).
CC NONHELICAL REGION 7 (NC7).
CC TRIPLE-HELICAL REGION 7 (COL7).
CC NONHELICAL REGION 8 (NC8).
CC TRIPLE-HELICAL REGION 8 (COL8).
CC NONHELICAL REGION 9 (NC9).
CC TRIPLE-HELICAL REGION 9 (COL9).
CC NONHELICAL REGION 10 (NC10).
CC TRIPLE-HELICAL REGION 10 (COL10).
CC NONHELICAL REGION 11 (NC11).
CC TRIPLE-HELICAL REGION 11 (COL11).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CELL ATTACHMENT SITE (POTENTIAL).
CC MISSING (IN SHORT ISOFORM).
CC AVPTQLPPQSNLQAPLGRSPAPDF -> MAPRWHLDLVL

```

FT CONFLICT 900 900 TSVLTLVAVSVNAE (IN SHORT ISOFORM).
FT CONFLICT 947 947 P -> L (IN REF. 4).
FT CONFLICT 964 964 A -> R (IN REF. 4).
FT CONFLICT 1157 1157 R -> P (IN REF. 4).
FT CONFLICT 1266 1266 P -> L (IN REF. 4).
FT CONFLICT 1276 1276 L -> F (IN REF. 4).
FT CONFLICT 1437 1437 L -> V (IN REF. 4).
SQ SEQUENCE 1527 AA; 156008 MW; 9645045AF140B513 CRC64;

Query Match 87.1%; Score 778; DB 1; Length 1527;
Best Local Similarity 85.8%; Pred. No. 1,7e-69;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGRGIRGADFOCFQOARAVAGLAGTFRATLSRLDLYSVIRADRAVPIV 60
DB 1356 VALNTPLSGRGIRGADFOCFQOARAVAGLAGTFRATLSRLDLYSVIRADRAVPIV 1415
QY 61 NLKDELLFPPSWALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
DB 1416 NLKDEVLSPSMDSLFSGGQGLQPGARIFSPDGRDVLRLHPMPQKSVWHGSDPNGRRLTE 1475
QY 121 SYCTWRTAPSATGQASSLGGRLLGQASASCHAIYVLCIENSFMTA 169
DB 1476 SYCTWRTETTGATGQASSLGGRLLGQASASCHAIYVLCIENSFMTS 1524

RESULT 3
CAIE_HUMAN STANDARD; PRT; 1388 AA.
ID CAIE_HUMAN STANDARD; PRT; 1388 AA.
AC P39059;
DC 01-FEB-1995 (Ref. 31, Created)
DT 01-FEB-1995 (Ref. 31, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE Collagen alpha 1(XV) chain precursor.
GN COL15A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical cord;
RA MEDLINE=94148920; PubMed=8106446;
RA Kivirikko S., Heilmaki P., Rehn M.V., Honkanen N., Myers J.C.,
RA Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of human type XV collagen and
RT exon-intron organization in the 3' region of the corresponding
RT gene."
RL J. Biol. Chem. 269:4773-4779(1994).
RN [2]
RP SEQUENCE OF 1-565 FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=94140817; PubMed=8307960;
RA Murgaki Y., Abe N., Ninomiya Y., Olsen B.R., Ooshima A.;
RT "The human alpha 1(XV) collagen chain contains a large amino-terminal
RT non-triple helical domain with a tandem repeat structure and homology
RT to alpha 1(XVIII) collagen."
RL J. Biol. Chem. 269:4042-4046(1994).
RN [3]
RP SEQUENCE OF 544-1252 FROM N.A.
RA MEDLINE=9306196; PubMed=1279671;
RA Myers J.C., Kivirikko S., Gordon M.K., Dion A.S., Pihlajaniemi T.;
RT "Identification of a previously unknown human collagen chain, alpha
RT 1(XV), characterized by extensive interruptions in the triple-helical
RT region."
RL Proc. Natl. Acad. Sci. U.S.A. 89:10144-10148(1992).
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN INTERNAL ORGANS
CC SUCH AS ADRENAL GLAND, PANCREAS AND KIDNEY.
CC -1- PTM: PROLINS AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
CC INTERRUPTED HELICES (FACIT) FAMILY.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L25266; AAA58429.1; -
DR EMBL: D21230; BAA04762.1; -
DR EMBL: L01697; -, NOT_ANNOTATED_CDS.
DR HSSP: P39061; IKOE.
DR Genew: HGNC:2192; COL15A1.
DR MIM: 120325; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR003129; TSPN.
DR Pfam: PF01391; Collagen; 5.
DR Pfam: PF02210; TSPN; 1.
DR SMART: SM00282; LamG; 1.
DR SMART: SM00210; TSPN; 1.
DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 1388
FT DOMAIN 26 555
FT DOMAIN 556 573
FT DOMAIN 574 618
FT DOMAIN 619 732
FT DOMAIN 733 763
FT DOMAIN 764 798
FT DOMAIN 799 822
FT DOMAIN 823 867
FT DOMAIN 868 878
FT DOMAIN 879 949
FT DOMAIN 950 963
FT DOMAIN 964 1013
FT DOMAIN 1014 1027
FT DOMAIN 1028 1045
FT DOMAIN 1046 1052
FT DOMAIN 1053 1107
FT DOMAIN 1108 1117
FT DOMAIN 1118 1132
FT DOMAIN 1133 1388
FT DOMAIN 1358 1388
FT REPEAT 358 408
FT REPEAT 409 459
FT REPEAT 460 509
FT REPEAT 510 555
FT CARBOHYD 306 306
FT CARBOHYD 324 324
FT CARBOHYD 687 687
FT CARBOHYD 807 807
FT CARBOHYD 814 814
FT CARBOHYD 1046 1046
FT CONFLICT 10 10
FT CONFLICT 49 49
FT CONFLICT 95 95
FT CONFLICT 150 150
FT CONFLICT 204 204
FT CONFLICT 409 409
SQ SEQUENCE 1388 AA; 141930 MW; 60822AD925A3093D CRC64;

Query Match 56.4%; Score 504; DB 1; Length 1388;
Best Local Similarity 56.9%; Pred. No. 3.2e-42;
Matches 95; Conservative 27; Mismatches 41; Indels 4; Gaps 1;

QY 2 ALNSPLSGRGIRGADFOCFQOARAVAGLAGTFRATLSRLDLYSVIRADRAVPIV 61
DB 1222 ALNPPSPGDR-----ADFOCFQOARAGLSTYRAFLSSRLDLYSVIRADRAVPIV 1277
QY 62 LKDELLFPPSWALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNGRRLTES 121

```

Db 1278 LKQGVFNWMDSTFSGGGQFNHPIFYSPDGRDITDPSPQKVIWHGSSPHGVRLVDN 1337
 QY 122 YCETWTEAPSATGQASLLGRLGSAASCHAAVIVLCIENSFMT 168
 Db 1338 YCQAMTADPAVAVGLASPLSTGKILQDKAVSCANRLIVLCIENSFMT 1384

RESULT 4

GLI_CHICK STANDARD; PRT; 556 AA.

AC P55378;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE zinc finger protein GLI1 (GLI) (Fragment).
 GN GLI1 OR GLI.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxId=9031;
 RN [1]
 RP MEDLINE=97105842; PubMed=8948590;
 RA Marigo V., Johnson R.L., Vortkamp A., Tabin C.J.;
 RT "Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development.";
 RL Dev. Biol. 180:273-283(1996).
 CC - FUNCTION: MAY REGULATE THE TRANSCRIPTION OF SPECIFIC GENES DURING NORMAL DEVELOPMENT. MAY PLAY A ROLE IN CRANIOFACIAL DEVELOPMENT AND DIGITAL DEVELOPMENT, AS WELL AS DEVELOPMENT OF THE CENTRAL NERVOUS SYSTEM AND GASTROINTESTINAL TRACT. IMPLICATED IN THE TRANSDUCTION OF SHH SIGNAL (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: Nuclear.
 CC - SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U60762; AAB51659.1; -.
 DR HSBP; P08151; 2GL1.
 DR InterPro; IPR000822; ZnF_C2H2.
 DR Pfam; PF00096; zf_C2H2; 5.
 DR SMART; SMO0355; ZnF_C2H2; 5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 5.
 KW zinc finger; Metal-binding; DNA-binding; Transcription regulation; Nuclear protein; Repeat.
 KM DOMAIN 247 399
 FT ZN FING 247 272 ZINC FINGERS.
 FT ZN FING 280 307 C2H2-TYPE.
 FT ZN FING 313 337 C2H2-TYPE.
 FT ZN FING 343 368 C2H2-TYPE.
 FT ZN FING 374 399 C2H2-TYPE.
 FT NON_TER 556 556
 SQ SEQUENCE 556 AA; 60215 MW; 722D2A5A1CA4D98 CRC64;

Query Match 8.5%; Score 76; DB 1; Length 556;
 Best Local Similarity 23.2%; Pred. No. 5.5;
 Matches 48; Conservative 21; Mismatches 66; Indels 72; Gaps 10;

QY 6 PLSG---GKRGIGADFO-C-----FOQARAVGLAG 32
 Db 20 PLHGASAGTGLGLDPRVCHQPNRLASHHGVLPTCTHPGAAADSRSTRGAGKLG 79
 QY 33 TFPAPF---LSSRLQDYSIVRRADRAAVPIVN-----LKDILLFPS----- 70

Db 80 KKRALSIPLSSSVLDQTVIRTPNSLVAFNRCASAGSYGHLSISTISLGYQNP 139
 QY 71 -----WEALFSGSEGPLRPGARIFSPDGKDYLRHTWPQKSVWHGSDPNGR-----LT 119
 Db 140 PQGQKQGQGLFSHTP-PLPPCSSHETLSSRPLGHLHTPARGRTIKKCOQLKERSLSSPLT 198

QY 120 ESYCETWTE---APSATGQASSLIG 142
 Db 199 AKYPEE-KSEGDISPASTGTODPLIG 224

RESULT 5

Y4BG_RHISN STANDARD; PRT; 271 AA.

AC P55374;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE Hypothetical protein Y4BG precursor.
 GN Y4BG.
 OS Rhizobium sp. (strain NGR234).
 OC Plasmid sym pNGR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OC NCBI_TaxId=394;
 RN [1]
 RP MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";
 RL Nature 387:394-401(1997).
 CC - SIMILARITY: NONE OBVIOUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB000066; AAB91622.1; -.
 KW Hypothetical protein; Plasmid; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 271
 FT CHAIN 271 271
 SQ SEQUENCE 271 AA; 30034 MW; 5572CF1461FCA63 CRC64;

Query Match 8.4%; Score 75; DB 1; Length 271;
 Best Local Similarity 21.1%; Pred. No. 3;
 Matches 28; Conservative 26; Mismatches 35; Indels 44; Gaps 7;

QY 4 NSPLSGMGKRGIRGADFOCFQARAVGLACTFPAPLSSRLQDLYSVRA-----DRAAV 57
 Db 174 NDAIATITKSLGADPEARQAITGTGSGEFS-----DDY-ISRCTPHMCRQEAL 225
 QY 58 PLYNLDDELFPSEWALFSGSEGPLKPGARIFSPDGKDYLRH-----TWQKSVWHGSDPR 114
 Db 226 LFLSADRDRAVAAW-----KP-----HKKIIVHPVQWPEKA----- 259
 QY 115 GRRLTESYCEWR 127
 Db 260 -KQELRAMAETWK 271

RESULT 6
 CRK_XENLA STANDARD; PRT; 296 AA.

ID CRK_XENLA
 AC P87378;
 DT 15-JUL-1998 (rel. 36, Created)
 DT 15-JUL-1998 (rel. 36, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE HYPOHETICAL lipoprotein yfm precursor.
OS YFM OR B2520.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCBI_TaxID=562;
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G., Ili, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
(potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, AE00338; AAC7573.1; -.
CC DR EcoGene; EG13394; yfm.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Hypothetical protein; Membrane; Lipoprotein; signal; Coiled coil;
KW Complete proteome.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 1653 HYPOTHETICAL LIPOPROTEIN YFM.
FT DOMAIN 1559 1589 COILED COIL (POTENTIAL).
FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 1653 AA; 181584 MW; 13109BC5CDB81AA CRC64;

Query Match 8.2%; Score 73.5; DB 1; Length 1653;
Best Local Similarity 24.4%; Pred. No. 35;
Matches 32; Conservative 21; Mismatches 47; Indels 31; Gaps 5;

QY 29 GLGTPAPAPSSRLQDY-----SYVRADRAVPVYINLDELFPSEWALFSGSE 79
DB 380 GADGYSKQFPMFGRDLYRGEETVILNGLADADGKALPNOPKLDVYKPGQVLRSVS 439
QY 80 GPLKPGARIFSFDGKDVLRHPTWPKS-----VWH---GSPDNGRRLTESCETWTE-- 129
DB 440 QP-----ENGLYHFTWLDGNAATGMHIRANGDNGYRMWDPRVDEPFMERM 487
QY 130 APARTGQASL 140
DB 488 ALNLTGKPTL 498

RESULT 9
PPX_ECOLI STANDARD; PRT; 512 AA.
AC P29014; P76981;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exopolysphatase (EC 3.6.1.11) (Exopolysphatase) (Metaphosphatase).
GN PPX OR B2502 OR Z3765 OR ECS3364.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCBI_TaxID=562, 83334;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=K12;
RX MEDLINE=93107072; PubMed=8380170;
RA Akiyama M., Crooke B., Kornberg A.;
RT "An exopolysphatase of Escherichia coli. The enzyme and its ppx
RT gene in a polysphatase operon."
RL J. Biol. Chem. 268:633-639(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G., Ili, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=9734980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Iseno K.,
Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
Mizouchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
Oshima T., Oyama S., Saito N., Stampel G., Satoh Y., Sivasubram S.,
Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
Yamagata S., Horichi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features."
RL DNA Res. 4:91-113(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., Ili, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Roesbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Portomais K.,
Apodaca J., Anantharaman T.S., Lin J., Yen J., Yen G., Schwartz D.C.,
Welch R.A., Blatner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [6]
RP SIMILARITY TO GPPA.
RX MEDLINE=94025037; PubMed=8212131;
RA Reitzer J., Reizer A., Sailer M.H., Jr., Bork B., Sander C.;
RT "Exopolysphatase phosphatase and guanosine pentaphosphate
RT phosphatase belong to the sugar kinase/actin/hsp 70 superfamily."
RL Trends Biochem. Sci. 18:247-248(1993).
CC -1- FUNCTION: DEGRADATION OF INORGANIC POLYPHOSPHATES. ORTHOPHOSPHATE
CC IS RELEASED PROGRESSIVELY FROM THE ENDS OF POLYPHOSPHATE OF CIRCA
CC 500 RESIDUES LONG. BUT CHAINS OF CIRCA 15 RESIDUES COMPLETE POORLY
CC WITH POLYPHOSPHATE AS SUBSTRATE.
CC -1- CATALYTIC ACTIVITY: {Polyphosphate} (n) + H(2)O =
CC {polyphosphate} (n-1) + phosphate.
CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC -1- SUBUNIT. HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: BELONGS TO THE GPPA / PPX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L06129; AAA24415.1; -
 DR EMBL; AE000336; AAC75555.1; -
 DR EMBL; D90878; BAA16390.1; -
 DR EMBL; D90880; BAA16392.1; -
 DR EMBL; AE005479; AAC57612.1; -
 DR EMBL; AP002561; BAB36787.1; -
 DR PIR; A45333; A45333.
 DR EcoGene; EG11403; Ppx.
 DR InterPro; IPR003695; Ppx GppA.
 DR Pfam; PF02541; Ppx-GppA; 1.
 KM HydroLase; Magnesium; Membrane; Complete proteome.
 FT INIT MET 0
 SQ SEQUENCE 512 AA; 58004 MW; 48611AF5D9FB9C3 CRC64;

Query Match 8.2%; Score 73; DB 1; Length 512;
 Best Local Similarity 21.9%; Pred. No. 10;
 Matches 46; Conservative 28; Mismatches 66; Indels 70; Gaps 11;

QY 22 FQAR-----AVGLAGTFR-----FLSSRLDLY 46
 DB 186 FQARMAAQKLTETWQFRIQGNVAMGASGTIKAAHEVLMENGEKGIITPRLELV 245
 QY 47 -SYRADRAAIVNLKDE--LFPSEALFSG-----SEGLKPGARIFS 91
 DB 246 KEVLRHNFASLSLPGESEKRTVFPGLALICGFDALARELRDGLAREGV-LYEM 304
 QY 92 DGR---DVRHPFWPKQSVHGSDFPGRRL---TESYCTWRTEAPS-ATGQASSLLG- 142
 DB 305 EGRHRHDVSRKTSLSANQNHIDSEQARRVLDTTMQMTEQMRQPKLAHPQLALLRW 364

QY 143 -----GRLLGQSAASCHNAVIVLCIENS 165
 DB 365 AAMLEHVEGLNINHSGLRHSAVI---LQNS 391

RESULT 10
 ID YP67 MYCTU STANDARD; PRT; 884 AA.
 AC 050654; 050731;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypochemical protein RV2567.
 GN RV2567 OR MT2643 OR MTCY227.34C OR MTCY9C4.01C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=8929587; PubMed=9634230;
 RA Cole S.T., Broese R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornby T., Osborne K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Bomojaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gall J., Mikula A.,
 RA Bishai W.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: SOME, TO M.TUBERCULOSIS RV2411C AND SYNECHOCYSTIS PCC
 CC 6803 SLL0335.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z77250; CAB01053.1; -
 DR EMBL; AE007098; AAK46956.1; -
 DR TIGR; MT2643; -
 DR TubercuList; RV2567; -
 KM Hypochemical protein; Complete proteome.
 FT CONFLICT 645 645 Q -> R (IN REF. 2).
 SQ SEQUENCE 884 AA; 95448 MW; 95D23A4D2DEB365 CRC64;

Query Match 8.1%; Score 72.5; DB 1; Length 884;
 Best Local Similarity 29.5%; Pred. No. 21;
 Matches 31; Conservative 15; Mismatches 26; Indels 33; Gaps 6;

QY 51 RADRAAIVNLKDELLFSPSEALFSGSEGLKPGARIFSFDG-----KQVLRHPFW 102
 DB 609 RADWIVA-----PS--TLWSLTVPDRPSPVSGVEGIALAAQAVRQLSNDTW 656

QY 103 -----PKSYWHSDFPGRRLTESYCTWRTEAPSATGQASSLLG 142
 DB 657 MWLANERAVEHRSDF-----PQSLAE--ADAVLSAQKETLAG 693

RESULT 11
 ID CRK_HUMAN STANDARD; PRT; 304 AA.
 AC P46108;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proto-oncogene C-crk (P38) (Adaptor molecule crk).
 GN CRK.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic lung; and placenta;
 RX MEDLINE=92334347; PubMed=1630456;
 RA Matsuda M., Tanaka S., Nagata S., Kojima A., Kurata T., Shibuya M.,
 RT "Two species of human CRK cDNA encode proteins with distinct
 RL biological activities."
 RL Mol. Cell. Biol. 12:3482-3489(1992).
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93390962; PubMed=8378094;
 RA Fiorotto T., Heisterkamp N., Groffen J., Benjes S., Morris C.;
 RT "CRK proto-oncogene maps to human chromosome band 17p13."
 RL Oncogene 8:2853-2855(1993).
 CC -1- FUNCTION: THE CRK-I AND CRK-II FORMS DIFFER IN THEIR BIOLOGICAL
 CC ACTIVITIES. CRK-II HAS LESS TRANSFORMING ACTIVITY THAN CRK-I. BOTH
 CC CRK-I AND CRK-II BIND TO MANY OF TYROSINE-PHOSPHORYLATED PROTEINS
 CC THAT BIND TO GRB2
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, CRK-I AND CRK-II (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE CRK-I FORM LACKS THE

```

CC LAST 100 RESIDUES.
CC -1- DOMAIN: THE C-TERMINAL SH3 DOMAIN FUNCTION AS A NEGATIVE MODULATOR
CC FOR TRANSFORMATION AND THE N-TERMINAL SH3 DOMAIN APPEARS TO
CC FUNCTION AS A POSITIVE REGULATOR FOR TRANSFORMATION.
CC -1- PTM: PHOSPHORYLATION OF CRK-II (40 kDa) GIVES RISE TO A 42 kDa
CC FORM.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, D10656; BAA01505.1; -.
CC EMBL, S65701; AAB28213.1; -.
CC HSSP; Q64010; 1CKA.
CC SWISS-2DPAGE; P46108; HUMAN.
CC Genew; HGNC:2362; CRK.
CC MIM: 164762; -.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00452; SH3DOMAIN.
CC ProDom; PD000066; SH3; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 2.
CC SMART; PS50001; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROTO-ONCOGENE; SH2 domain; SH3 domain; Repeat; Alternative splicing;
CC PHOSPHORYLATION.
CC KW DOMAIN 13 118 SH2.
CC FT DOMAIN 132 192 SH3 1.
CC FT DOMAIN 256 296 SH3 2.
CC FT VARSPPLIC 205 304 MISSING (IN ISOFORM CRK-I).
CC SQ SEQUENCE 304 AA; 33872 MW; D74A83ED1PFC0BEC CRC64;

Query Match 8.1%; Score 72; DB 1; Length 304;
Best Local Similarity 27.8%; Pred. No. 6.9;
Matches 20; Conservative 12; Mismatches 24; Indels 16; Gaps 3;

QY 88 IEFPGDK-----DVLPHPTPKQSVVHGSDPNRR--LTESYCEFWTEAPSGTG 135
DB 140 LFPDNGNDEEDLPFKKQDILIRIDKPEEQWMADESEKGMIPVPEVKYR-----PASA 195
QY 136 QASLLGGRLLG 147
DB 196 SVSALLIGNOEG 207

RESULT 12
CRK_MOUSE STANDARD; PRT; 304 AA.
AC Q64010;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE PROTO-ONCOGENE C-CRK (P38) (Adapter molecule crk).
GN CRK OR CRKO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94239744; PubMed=8183562;

```

```

RA Ogawa S., Toyoshima H., Kozutsumi H., Hagihara K., Sakai R.,
RA Tanaka T., Hirano N., Mano H., Yazaki Y., Hirai H.;
RA "The C-terminal SH3 domain of the mouse C-Crk protein negatively
RA regulates tyrosine-phosphorylation of Crk associated p130 in rat 3Y1
RA cells.";
RA Oncogene 9:1669-1678(1994).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 134-190.
RX MEDLINE=95253821; PubMed=7735837;
RA Wu X., Knudsen B., Feller S.M., Zheng J., Sali A., Cowburn D.,
RA Hanafusa H., Kurlyan J.;
RA "Structural basis for the specific interaction of lysine-containing
RA proline-rich peptides with the N-terminal SH3 domain of C-Crk.";
RA Structure 3:215-226(1995).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 133-191.
RX MEDLINE=99069628; PubMed=9851931;
RA Nguyen J.T., Turk C.W., Cohen P.E., Zuckermann R.N., Lim W.A.;
RA "Exploiting the basis of proline recognition by SH3 and WW domains:
RA design of N-substituted inhibitors.";
RA Science 282:2088-2092(1998).
CC -1- FUNCTION: THE CRK-I AND CRK-II FORMS DIFFER IN THEIR BIOLOGICAL
CC ACTIVITIES. CRK-II HAS LESS TRANSFORMING ACTIVITY THAN CRK-I. BOTH
CC CRK-I AND CRK-II BIND TO MANY OF TYROSINE-PHOSPHORYLATED PROTEINS
CC THAT BIND TO GRB2.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CRK-I AND CRK-II (SHOWN HERE);
CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE CRK-I FORM LACKS THE
CC LAST 100 RESIDUES.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- DOMAIN: THE C-TERMINAL SH3 DOMAIN FUNCTION AS A NEGATIVE MODULATOR
CC FOR TRANSFORMATION AND THE N-TERMINAL SH3 DOMAIN APPEARS TO
CC FUNCTION AS A POSITIVE REGULATOR FOR TRANSFORMATION (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S72408; AAB30755.1; -.
CC PDB; 1CKB; 08-MAY-95.
CC PDB; 1CKB; 08-MAY-95.
CC PDB; 1B07; 06-JAN-99.
CC MGD; MGJ:88508; Crko.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00452; SH3DOMAIN.
CC ProDom; PD000066; SH3; 1.
CC ProDom; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 2.
CC SMART; PS50001; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC PROTO-ONCOGENE; SH2 domain; SH3 domain; Repeat; Alternative splicing;
CC KW PHOSPHORYLATION; 3D-structure.
CC FT DOMAIN 13 118 SH2.
CC FT DOMAIN 132 192 SH3 1.
CC FT DOMAIN 256 296 SH3 2.
CC FT VARSPPLIC 205 304 MISSING (IN ISOFORM CRK-I).
CC SQ SEQUENCE 304 AA; 33814 MW; 5491896FC7A89065 CRC64;

Query Match 8.1%; Score 72; DB 1; Length 304;
Best Local Similarity 27.8%; Pred. No. 6.9;
Matches 20; Conservative 12; Mismatches 24; Indels 16; Gaps 3;

```

Qy 88 IFSPDGK-----DVLRHPTWPKSVHSGDPNGRR--LTESYCETWRTAPSATG 135
 Db 140 LFDPNNGDEEDLPFKKGDILIRDKPEEQWMAEDSGKGMIPVYVEKXR-----PASA 195
 Qy 136 QASLLGRLLG 147
 Db 196 SVSALIGNGNEG 207

RESULT 13

CRK_RAT ID CRK_RAT STANDARD; PRT; 304 AA.
 AC 063768;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proto-oncogene C-crk (P38) (Adapter molecule crk).
 GN CRK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=97057214; PubMed=8901553;
 RA Kizaka-Kondoh S., Matsuda M., Okayama H.;
 RT "Crkl signals from epidermal growth factor receptor to Ras."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12177-12182(1996).
 CC -1- FUNCTION: THE CRK-I AND CRK-II FORMS DIFFER IN THEIR BIOLOGICAL
 CC ACTIVITIES. CRK-II HAS LESS TRANSFORMING ACTIVITY THAN CRK-I. BOTH
 CC CRK-I AND CRK-II BIND TO MANY OF TYROSINE-PHOSPHORYLATED PROTEINS
 CC THAT BIND TO GRB2.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CRK-I AND CRK-II (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE CRK-I FORM LACKS THE
 CC LAST 100 RESIDUES.
 CC -1- TISSUE SPECIFICITY: CRK-II IS EXPRESSED IN ALL TISSUES AND CELLS
 CC WHEREAS CRK-I IS EXPRESSED AT LOWER LEVEL AND IN LIMITED CELL-
 CC TYPES.
 CC -1- DOMAIN: THE C-TERMINAL SH3 DOMAIN FUNCTION AS A NEGATIVE MODULATOR
 CC FOR TRANSFORMATION AND THE N-TERMINAL SH3 DOMAIN APPEARS TO
 CC FUNCTION AS A POSITIVE REGULATOR FOR TRANSFORMATION (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D44481; BAA07924.1; -.
 DR HSSP; Q64010; ICKA.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS00001; SH2; 1.
 DR PROSITE; PS00002; SH3; 1.
 KM Proto-oncogene; SH2 domain; SH3 domain; Repeat; Alternative splicing;
 KM Phosphorylation.
 FT DOMAIN 13 118 SH2.
 FT DOMAIN 132 192 SH3 1.
 FT DOMAIN 256 296 SH3 2.

FT VARSPLIC 205 304 MISSING (IN ISOFORM CRK-I).
 FT VARIANT 244 244 Q -> R (NRK-23 INACTIVE MUTANT).
 FT VARIANT 253 253 K -> E (NRK-23 INACTIVE MUTANT).
 SQ SEQUENCE 304 AA; 33844 MW; 4CFBFB5B872E265 CRC64;

Query Match

Best Local Similarity 27.8%; Pred. No. 6.9; DB 1; Length 304;
 Matches 20; Conservative 12; Mismatches 24; Indels 16; Gaps 3;

Qy 88 IFSPDGK-----DVLRHPTWPKSVHSGDPNGRR--LTESYCETWRTAPSATG 135
 Db 140 LFDPNNGDEEDLPFKKGDILIRDKPEEQWMAEDSGKGMIPVYVEKXR-----PASA 195
 Qy 136 QASLLGRLLG 147
 Db 196 SVSALIGNGNEG 207

RESULT 14

H575 CANAL ID H575 CANAL STANDARD; PRT; 613 AA.
 AC P87232;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Heat shock protein SSB1.
 GN SSB1 OR HSP70B.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Microsporite Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NO-1;
 RX MEDLINE=9734370; PubMed=9200817;
 RA Maneu V., Cervera A.M., Martinez J.P., Gosalbo D.;
 RT "Molecular cloning of a Candida albicans gene (SSB1) coding for a
 RT protein related to the Hsp70 family."
 RL Yeast 13:677-681(1997).
 CC -1- FUNCTION: MAY AID IN THE PASSAGE OF THE NASCENT POLYPEPTIDE CHAIN
 CC THROUGH THE RIBOSOME CHANNEL INTO THE CYTOSOL. SUCH AN INTERACTION
 CC COULD BE CRUCIAL FOR CONTINUOUS TRANSPORT OF THE POLYPEPTIDE.
 CC COULD SERVE TO PREVENT THE NASCENT POLYPEPTIDE FROM INTERFERING
 CC WITH TRANSLATION BY CLOGGING THE RIBOSOME CHANNEL (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH TRANSLATING
 CC RIBOSOMES; MAY BIND DIRECTLY TO THE NASCENT POLYPEPTIDE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X97723; CA66308.1; -.
 DR HSSP; P08107; 1HCO.
 DR COMPUTEYEAST-2DPAGE; P87222; -.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PR00301; HEATSHOCK70.
 DR ProDom; PD000089; HSP70; 1.
 DR ProDom; PD000089; HSP70; 1.
 DR PROSITE; PS00287; HSP70_1; 1.
 DR PROSITE; PS00329; HSP70_2; 1.
 DR PROSITE; PS01036; HSP70_3; 1.
 KM Heat shock; ATP-binding; Multigene family; Protein biosynthesis.
 SQ SEQUENCE 613 AA; 66432 MW; 5FBA9F8F9327F9 CRC64;

Query Match

Best Local Similarity 25.6%; Pred. No. 16;
 Matches 33; Conservative 20; Mismatches 50; Indels 26; Gaps 6;

QY 32 GTRAFALSSRLQDLYSIVRRADRAVPIVNLKDELLFSSWEALFS-----GSEGPL 82
 DB 4 GVFGAIGIDLGITYSCVATYDSAVEIIANEQGNRVTPSPVAFSEBRLIGDAKQNAL 63
 QY 83 KPGARIF-----SPDGKDVLRH-PTWPQKSWHSGSDNGRRRLTE-STCEWTREAPS 132
 DB 64 NPKTYFEDAKRLIGRAFDDESVOIKSWPKVY-----ESNGQPLLEVEYLDETKTFSPQ 119
 QY 133 ATGOASSLL 141
 DB 120 ---EISSMV 125
 RESULT 15
 STRN MOUSE STANDARD; PRT; 780 AA.
 AC 055106;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Striatin.
 GN STRN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxId=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Brain;
 RA Mogrich A., Mattei M.-G., Bartoli M., Rakitina T., Ballat G.,
 RA Monneron A., Casteels F.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP TISSUE SPECIFICITY.
 RA MEDLINE=20347911; PubMed=10748158;
 RA Casteels F., Rakitina T., Gaillard S., Mogrich A., Mattei M.-G.,
 RA Monneron A.;
 RT "Zinedin, SG2NA, and striatin are calmodulin-binding, WD repeat
 RT proteins principally expressed in the brain.";
 RL J. Biol. Chem. 275:19970-19977(2000).
 CC -1- FUNCTION: BINDS CALMODULIN IN A CALCIUM DEPENDENT MANNER. MAY
 CC FUNCTION AS SCAFFOLDING OR SIGNALING PROTEIN.
 CC -1- SUBUNIT: INTERACTS WITH PROTEIN PHOSPHATASE 2A (PP2A) (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-BOUND.
 CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN BRAIN BUT IS ALSO
 CC EXPRESSED AT LOW LEVELS IN VARIOUS TISSUES SUCH AS KIDNEY, SPLEEN,
 CC SKELETAL MUSCLE AND LUNG.
 CC -1- SIMILARITY: BELONGS TO THE STRIATIN FAMILY OF WD-REPEAT PROTEINS.
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ223777; CA11545.1; -.
 DR MGD; MG1:133757; Strn.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 6.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 3.
 DR SMART; SM00320; WD40; 5.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS00682; WD_REPEATS_2; 4.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Calmodulin-binding; Repeat; WD repeat; Coiled coil.
 FT DOMAIN 53 120 COILED COIL (POTENTIAL).
 FT DOMAIN 149 166 CALMODULIN-BINDING (POTENTIAL).
 FT REPEAT 461 500 WD 1.

FT REPEAT 514 553 WD 2.
 FT REPEAT 567 606 WD 3.
 FT REPEAT 662 701 WD 4.
 FT REPEAT 704 743 WD 5.
 FT REPEAT 750 779 WD 6.
 FT DOMAIN 37 45 POLY-ALA.
 FT SITE 55 63 CAVEOLIN-BINDING (POTENTIAL).
 SQ SEQUENCE 780 AA; 86013 MW; DBD1104FP95BC08 CRC64;
 Query Match 7.8%; Score 70; DB 1; Length 780;
 Best Local Similarity 23.3%; Pred. No. 32;
 Matches 50; Conservative 19; Mismatch 66; Indels 80; Gaps 12;
 QY 2 ALNSPLSGMRGIRGADFOCFQOARAVG-LAGTFRAFLSRLQDLYSIVRRADRAVPIV 60
 DB 398 ALTFPPSSGKSPFIMGAD-ELAESELGELIAG-----LTVANEDSLAYDIA 443
 QY 61 NLKDELLFPSWEALFSGSSEGPLKPGARIFSPDGKDVLR-HPTWP----- 103
 DB 444 NNKDALR-KTWNPKEF-----LRS-----HFDGIRALAFPIEPVLITASEDHTLKMWNL 492
 QY 104 -----QKSWHSGSDP-----NRRLL-----TESYCETWTEAPSA-- 133
 DB 493 QKTAAPAKKSTSLDVEPIYTFRAHKGPEVLCVWSSNGECYSGGTDRIOSWSTNPVDP 552
 QY 134 -TGOASSLLGRLIGQS-----AASCHAYIVLC 161
 DB 553 YDAVDPSTVLRGPLGHTGTDVWGLAYSAHQRLLSC 587

Search completed: February 11, 2003, 20:41:47
 Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:37:24 ; Search time 44 Seconds

(without alignments)
796.091 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893
Sequence: 1 VALNSPLSGMRGIRGADPQ.....ASCHAYIVLCIENSFWTAS 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	187	4 Q8WX15	Q8WX15 homo sapien
2	778	87.1	184	11 Q9JXK6	Q9JXK6 mus musculu
3	778	87.1	1140	11 Q61434	Q61434 mus musculu
4	778	87.1	1174	11 Q62001	Q62001 mus musculu
5	763	85.4	226	11 Q9QZD2	Q9QZD2 rattus norv
6	757	84.8	171	11 Q9WUW5	Q9WUW5 mus musculu
7	732	80.9	160	11 Q9CRT2	Q9CRT2 mus musculu
8	711	79.6	1344	13 Q93419	Q93419 gallus gall
9	654	73.2	1315	13 Q8QHL9	Q8QHL9 xenopus lae
10	529	59.2	102	4 Q96T70	Q96T70 homo sapien
11	515	57.7	1367	11 Q95206	Q95206 mus musculu
12	512	57.3	1367	11 Q9EAD9	Q9EAD9 mus musculu
13	505	56.6	1388	4 Q9Y4M4	Q9Y4M4 homo sapien
14	418.5	46.9	581	5 Q9VSO9	Q9VSO9 drosophila
15	357	40.0	650	5 Q17866	Q17866 caenorhabdi
16	357	40.0	778	5 Q9U9K6	Q9U9K6 caenorhabdi

17	357	40.0	1117	5 Q9U9K7	Q9U9K7 caenorhabdi
18	82.5	9.2	208	16 Q92K28	Q92K28 rhizobium m
19	82	9.2	498	16 Q9KXK2	Q9KXK2 streptomyce
20	80	9.0	493	16 Q98AT9	Q98AT9 rhizobium 1
21	80	9.0	651	5 Q9VPA9	Q9VPA9 drosophila
22	79	8.8	477	10 Q9SMY7	Q9SMY7 arabidopsis
23	79	8.8	525	10 Q94JL8	Q94JL8 arabidopsis
24	78.5	8.8	285	16 Q98EUS	Q98EUS rhizobium 1
25	78	8.7	1715	6 Q9GLM4	Q9GLM4 bos taurus
26	77.5	8.7	395	11 Q9QUP4	Q9QUP4 mus musculu
27	77	8.6	314	16 Q8U8N8	Q8U8N8 agrobacteri
28	76	8.5	204	4 Q96GA9	Q96GA9 homo sapien
29	76	8.5	636	4 Q60624	Q60624 homo sapien
30	75.5	8.5	334	11 Q9WTJ8	Q9WTJ8 mus musculu
31	75.5	8.5	346	5 Q9VNS3	Q9VNS3 drosophila
32	75.5	8.5	904	10 Q48541	Q48541 hordeum vul
33	75	8.4	351	16 Q99X21	Q99X21 stephilococ
34	74.5	8.3	850	4 Q14425	Q14425 homo sapien
35	74.5	8.3	962	10 Q9S7S8	Q9S7S8 hordeum vul
36	74	8.3	919	15 Q9WPP0	Q9WPP0 chimpanzee
37	73.5	8.2	325	4 Q9Y247	Q9Y247 homo sapien
38	73.5	8.2	695	10 Q8S975	Q8S975 oryza sativ
39	73.5	8.2	905	10 Q9PYV0	Q9PYV0 hordeum vul
40	73.5	8.2	958	5 Q8SX15	Q8SX15 drosophila
41	73.5	8.2	1612	5 Q9VYQ2	Q9VYQ2 drosophila
42	73.5	8.2	1653	16 Q8XA93	Q8XA93 escherichia
43	73	8.2	435	4 Q9UN18	Q9UN18 homo sapien
44	73	8.2	435	4 Q9UKD5	Q9UKD5 homo sapien
45	73	8.2	435	4 Q12794	Q12794 homo sapien

ALIGNMENTS

RESULT 1

ID	Q8WX15	PRELIMINARY;	PRT;	187 AA.
AC	Q8WX15;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DE	Collagen XVIII (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21409408; PubMed=11517600;			
RA	Feng Y., Cui L.B., Liu C.X., Ma Q.J.;			
RT	"Inhibition effect in vitro of purified endostatin expressed in Pichia			
RT	pastoris."			
RL	Sheng Wu Gong Cheng Xue Bao 17:278-282 (2001).			
DR	EMBL; AF165592; MAF37720.1; -.			
FT	NON TER			
FT	SEQUENCE 187 AA; 20448 MW; 72B1047D85838CD3 CRC64;			

Query Match 100.0%; Score 893; DB 4; Length 187;

Best Local Similarity 100.0%; Pred. No. 4.7e-83; Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	VALNSPLSGMRGIRGADPQFOARAVAGLAFRAFLSSRLDLYSVRRADAAVPIV 60
DB	17	VALNSPLSGMRGIRGADPQFOARAVAGLAFRAFLSSRLDLYSVRRADAAVPIV 76
QY	61	NLDELLFPSEWALFSGSEGLPLPGARIFSGDKDYLRLHPTWQKSVWQSDPGRRLTE 130
DB	77	NLDELLFPSEWALFSGSEGLPLPGARIFSGDKDYLRLHPTWQKSVWQSDPGRRLTE 136
QY	121	SYETRTREPAATGQASSILGRLIGQSAASCHAYIVLCIENSFWTAS 170
DB	137	SYETRTREPAATGQASSILGRLIGQSAASCHAYIVLCIENSFWTAS 186

RESULT 2
09UK63
ID 09UK63 PRELIMINARY; PRT; 184 AA.
AC 09UK63;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Endostatin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHINESE KUNMING;
RA Jia S., Zhu F., Xing G., Yu Y., Duan C., Xiu R.-J., He F.;
RT "Anticancer treatment of targeted fusion protein delivery to tumor neovascuclature";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257775; AAF69009.1; -.
DR HSSP; P39061; IKOE.
FT NON_TER 1 1
FT NON_TER 184 184
SQ SEQUENCE 184 AA; 20376 MW; AC06F9DBD103412A CRC64;
Query Match 87.1%; Score 778; DB 11; Length 184;
Best Local Similarity 85.8%; Pred. No. 2,4e-71;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADPFCFOQARAVGLAGTFRAFLSSRLDLYSIYRADAAPVIV 60
DB 13 VALNTPLSGMGRGIRGADPFCFOQARAVGLAGTFRAFLSSRLDLYSIYRADAAPVIV 72
QY 61 NLKDELLFPSEWALFSGSEGPLKRGARIFSPDGDVLRHPTWPKSVHSGSDPNGRRLTE 120
DB 73 NLKDEVSPSWDSLFSSGQQLPGARIFSPDGRDVLRHPTWPKSVHSGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLGRLGQASCHHAYIVLCIENSPMTA 169
DB 133 SYCETWRTETTGATGQASSLSGRLLGQKASCHNSYIVLCIENSPMTS 181

RESULT 3
061434
ID 061434 PRELIMINARY; PRT; 1140 AA.
AC 061434;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Collagen (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Abe N., Muragaki Y., Yoshioka H., Inoue H., Ninomiya Y.;
RT "Identification of a novel collagen chain represented by extensive RT interruptions in the triple-helical region.";
RL Cell. Mol. Biol. Res. 196:576-582(1993).
DR EMBL; D17546; BAA04483.1; -.
DR HSSP; P39061; IKOE.
DR MGD; MGI:88449; Col15a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
FT NON_TER 1 1
FT NON_TER 1140 1140
SQ SEQUENCE 1140 AA; 115156 MW; 8B0C7E68623BDFF CRC64;

Query Match 87.1%; Score 778; DB 11; Length 1140;
Best Local Similarity 85.8%; Pred. No. 2,4e-70;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADPFCFOQARAVGLAGTFRAFLSSRLDLYSIYRADAAPVIV 60
DB 969 VALNTPLSGMGRGIRGADPFCFOQARAVGLAGTFRAFLSSRLDLYSIYRADAAPVIV 1028
QY 61 NLKDELLFPSEWALFSGSEGPLKRGARIFSPDGDVLRHPTWPKSVHSGSDPNGRRLTE 120
DB 1029 NLKDEVSPSWDSLFSSGQQLPGARIFSPDGRDVLRHPTWPKSVHSGSDPNGRRLTE 1088
QY 121 SYCETWRTAPSATGQASSLGRLGQASCHHAYIVLCIENSPMTA 169
DB 1089 SYCETWRTETTGATGQASSLSGRLLGQKASCHNSYIVLCIENSPMTS 1137

RESULT 4
062001
ID 062001 PRELIMINARY; PRT; 1774 AA.
AC 062001; 060672;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Procollagen, type XVIII, alpha 1 precursor (XVIII) collagen (Procollagen, type XVIII, alpha 1) (Alpha-1 type XVIII collagen).
GN COL18A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY TAIL CULTURE;
RX MEDLINE=94245707; PubMed=8188673;
RA Rehn M., Hintikka E., Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen chain.";
RL J. Biol. Chem. 269:13929-13935(1994).
RN [2]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=94240112; PubMed=8183894;
RA Rehn M., Pihlajaniemi T.;
RT "Alpha 1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).
RN [3]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=95181468; PubMed=7876242;
RA Rehn M., Pihlajaniemi T.;
RT "Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. The longest form contains a novel motif homologous to rat U. Biol. Chem. 270:4705-4711(1995).
RL J. Biol. Chem. 270:4705-4711(1995).
DR EMBL; U03715; AAC52903.1; -.
DR EMBL; U03716; AAC52903.1; JOINED.
DR EMBL; U03718; AAC52903.1; JOINED.
DR EMBL; U34607; AAC52903.1; JOINED.
DR EMBL; U34608; AAC52903.1; JOINED.
DR EMBL; U34609; AAC52903.1; JOINED.
DR EMBL; U34610; AAC52903.1; JOINED.
DR EMBL; U34611; AAC52903.1; JOINED.
DR EMBL; U34612; AAC52903.1; JOINED.
DR EMBL; U34613; AAC52903.1; JOINED.
DR EMBL; U34613; AAC52903.1; JOINED.
DR EMBL; U1637; AAC52179.1; -.
DR HSSP; P39061; IKOE.
DR MGD; MGI:88451; Col18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000024; Fz_domain.

DR Interfero; IPR001791; Laminin_G.
 DR Pfam; PFO1391; Collagen; 8.
 DR Pfam; PFO1392; Fz; 1.
 DR Pfam; PFO2210; TSPN; 1.
 DR Prodom; PD000007; Collagen; 1.
 DR SMART; SM00063; FRL; 1.
 DR SMART; SM00282; LamiG; 1.
 DR SMART; SM00210; TSPN; 1.
 DR PROSITE; PS00038; FZ; 1.
 KM Signal.
 SQ SEQUENCE 1774 AA; 182229 MW; CF4D9BC9E8BEF232 CRC64;

Query Match 87.1%; Score 778; DB 11; Length 1774;
 Best Local Similarity 85.8%; Pred. No. 4.2e-70;
 Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGRADFOCFQOARAVGLAGTFRAFLSSRLDLYSIVRRADRAAVPIY 60
 DB 1603 VALNTPLSGMGRIGRADFOCFQOARAVGLSGTFRAFLSSRLDLYSIVRRADRGVPIY 1662
 QY 61 NLKDELLFPGMEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDPGRRLTE 120
 DB 1663 NLKDEVLPSPWDLTFSGSGQLPGARIFSPDGRDVLRRHPAWPKSVWHSDDPSGRRLME 1722
 QY 121 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHHAYIVLCIENSFMTA 169
 DB 1723 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHNSYIVLCIENSFMTS 1771

RESULT 5
 Q9QZD2 PRELIMINARY; PRT; 226 AA.
 ID Q9QZD2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Collagen XVIII (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20227226; PubMed=1076159;
 RA Perletti G., Concarì P., Giardini R., Marras E., Piccinini F.,
 RA Folkmann J., Chen L.,
 RT "Antitumor activity of endostatin against carcinogen-induced rat
 RT primary mammary tumors";
 RL Cancer Res. 60:1793-1796(2000).
 DR EMBL; AF189709; AAF00975.1; -.
 DR HSSP; P39061; 1KOE.
 FT NON_TER 1
 SQ SEQUENCE 226 AA; 25350 MW; 38B83C0486C0E949 CRC64;

Query Match 85.4%; Score 763; DB 11; Length 226;
 Best Local Similarity 84.6%; Pred. No. 1e-69;
 Matches 143; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGRADFOCFQOARAVGLAGTFRAFLSSRLDLYSIVRRADRAAVPIY 60
 DB 55 VALNTPLSGMGRIGRADFOCFQOARAVGLSGTFRAFLSSRLDLYSIVRRADRSSVPIY 114
 QY 61 NLKDELLFPGMEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDPGRRLTE 120
 DB 115 NLKDEVLPSPWDLTFSGSGQLHSGARIFSPDGRDVLRRHPAWPKSVWHSDDPSGRRLME 174
 QY 121 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHHAYIVLCIENSFMTA 169
 DB 175 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHNSYIVLCIENSFMTS 223

RESULT 6
 Q9WUW5 PRELIMINARY; PRT; 171 AA.
 ID Q9WUW5;
 AC Q9WUW5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Collagen type XVIII, alpha (I) chain (Fragment).
 GN COL18A1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Jia J.D., Bauer M., Eberspaecher U., Donner P., Schuppan D.,
 RT "Partial 3'-terminal cDNA sequence of rat collagen XVIII/endostatin.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Jia J.D., Bauer M., Sedlaczek N., Rühl M., Riecken E.O., Schuppan D.,
 RT "Temporal expression of collagen XVIII/endostatin in acute and
 RT chronic liver injuries";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ236873; CAB44263.1; -.
 DR HSSP; P39061; 1KOE.
 FT NON_TER 1
 FT NON_TER 171
 SQ SEQUENCE 171 AA; 18933 MW; 81BE2EE3FC2C8E72 CRC64;

Query Match 84.8%; Score 757; DB 11; Length 171;
 Best Local Similarity 80.0%; Pred. No. 2.9e-69;
 Matches 142; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGRADFOCFQOARAVGLAGTFRAFLSSRLDLYSIVRRADRAAVPIY 60
 DB 5 VALNTPLSGMGRIGRADFOCFQOARAVGLSGTFRAFLSSRLDLYSIVRRADRSSVPIY 64
 QY 61 NLKDELLFPGMEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDPGRRLTE 120
 DB 65 NLKDEVLPSPWDLTFSGSGQLHSGARIFSPDGRDVLRRHPAWPKSVWHSDDPSGRRLME 124
 QY 121 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHHAYIVLCIENSFMTA 167
 DB 125 SYCEWRTETATGATGQASSLSLGRLLQGSAASCHNSYIVLCIENSFMTS 171

RESULT 7
 Q9CRT2 PRELIMINARY; PRT; 160 AA.
 ID Q9CRT2;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Procollagen, type XVIII, alpha 1 (Fragment).
 GN COL18A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OK NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawaki J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi T., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Ozaki Y., Gotohori T., Bono H., Kaubawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Betsalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsumoto Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schriml L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Maehio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G., Blake J., Boffelli D., Botunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., RA Guetlinck S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Maehima J., Mazzarelli U., Mombauts P., RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmink L., RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014292; BAB29249.1; -.
DR HSSP; P39061; 1KOE.
DR MGD; MGI:88451; Coll18a1.
FT NON TER 1
SQ SEQUENCE 160 AA; 17725 MW; 60F853D77C375D2 CRC64;

Query Match 80.9%; Score 722; DB 11; Length 160;
Best Local Similarity 85.4%; Pred. No. 9,9e-66;
Matches 134; Conservative 12; Mismatches 11; Indels 0; Gaps 0;
QY 13 GIRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRAAVPIYVNLKDELLFSPSWE 72
DB 1 GIRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRGVPIYVNLKDEVLSPSWD 60
QY 73 ALFSGSGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDNGSRRLTESYCEMTREAPS 132
DB 61 SLFSGSGQQLDPGARIFSPDRDVLRRHPWPKSVHMGSDPSGRRLMESYCEMTRETTG 120
QY 133 ATGQASSLLGRLGQSAASCHAYIVLCIENSFMTA 169
DB 121 ATGQASSLLGRLLEQKASCHNYIVLCIENSFMTS 157

RESULT 8
Q93419 PRELIMINARY; PRT; 1344 AA.
AC Q93419;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DR Collagen XVIII precursor.
OS Gallus gallus (Chicken).
OC Birkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98411346; PubMed=9738008;
RA Halfter W., Dong S., Schurer B., Cole G.J.;
RT "Collagen XVIII is a basement membrane heparan sulfate proteoglycan.";
RN J. Biol. Chem. 273:25404-25412(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Halfter W., Dong S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083440; AAC33294.2; -.
DR HSSP; P39061; 1KOE.
DR InterPro; IPR000087; Coll18n.
DR InterPro; IPR001791; Lam18n_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 2.
DR SMART; SM00282; Lamg; 1.
DR SMART; SM00210; TSPN; 1.
KM Signal
FT SIGNAL 1 26 POTENTIAL.
SQ SEQUENCE 1344 AA; 137402 MW; 7AA36B4FE940CCD CRC64;

Query Match 79.6%; Score 711; DB 13; Length 1344;
Best Local Similarity 77.6%; Pred. No. 2e-63;
Matches 132; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
QY 1 VALNSPLSGMRIGRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRAAVPIY 60
DB 1173 VALNSPLSGMRIGRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRAAVPIY 1232
QY 61 NLKDELLFSPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDNGSRRLTE 120
DB 1233 NLKDEVLFPSWEALFSGSEGPLKPGARILISFDGRDILQDSAWPQKSIWMSGDAKGRRLPE 1292
QY 121 SYCEWTRTEAPSATGQASSLLGRLGQSAASCHAYIVLCIENSFMTS 170
DB 1293 SYCEWTRTEAPSATGQASSLLGRLGQSAASCHAYIVLCIENSFMTA 1342

RESULT 9
Q8QHL9 PRELIMINARY; PRT; 1315 AA.
AC Q8QHL9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Type XVIII collagen alpha1 chain.
OS Xenopus laevis (African clawed frog).
OC Birkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishino T., Sekimizu K., Natori S., Kubo T.;
RT "Identification and characterization of genes expressed selectively in the regenerating tail of *Xenopus laevis* tadpole.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047066; BAB84674.1; -.
KW Collagen.
SQ SEQUENCE 1315 AA; 134946 MW; 0C56C235DE058365 CRC64;

Query Match 73.2%; Score 654; DB 13; Length 1315;
Best Local Similarity 73.2%; Pred. No. 1.2e-57;
Matches 123; Conservative 16; Mismatches 29; Indels 0; Gaps 0;
QY 1 VALNSPLSGMRIGRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRAAVPIY 60
DB 1144 VALNSPLSGMRIGRADPQCFOQARAVGLAGTFRAFLSSRLDLYSIYRRADRAAVPIY 1203
QY 61 NLKDELLFSPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDNGSRRLTE 120
DB 1204 NLKDEVLYDNWESLFGSESEQMPGARIFSPDGKDVLRHPTWPKSVHMGSDAKGRRLTE 1263
QY 121 SYCEWTRTEAPSATGQASSLLGRLGQSAASCHAYIVLCIENSFMT 168
DB 1264 SYCEWTRTEAPSATGQASSLLGRLGQSAASCHAYIVLCIENSFMT 1311

RESULT 10
Q96T70 PRELIMINARY; PRT; 102 AA.
AC Q96T70;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Endostatin variant (Fragment).
OS Homo sapiens (Human).
OC Birkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Deininger M.H., Trautmann K., Schluessener H.J.;
RT "Endostatin promotes delayed secondary damage following traumatic

```
RT brain injury."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333247; AAK50626.1; -.
FT NON TER 1
FT NON TER 102
SQ SEQUENCE 102 AA; 11147 MW; ECAC47AA6420947D CRC64;

Query Match
Best Local Similarity 59.2%; Score 529; DB 4; Length 102;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 59 IYNLKELLPSWEALFSGSEGLPKRGARIFSPDGKDVLRHPTWPKSVHSGSDPNRRLL 118
DB 1 IYNLKELLPSWEALFSGSEGLPKRGARIFSPDGKDVLRHPTWPKSVHSGSDPNRRLL 60

QY 119 TESYCETWRTSPATGQASLLGGRLLGQSAASCHHAYIVL 160
DB 61 TESYCETWRTSPATGQASLLGGRLLGQSAASCHHAYIVL 102

RESULT 11
ID 035206 PRELIMINARY; PRT; 1367 AA.
AC 035206;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Type XV collagen.
GN COL15A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
[1]
RX SEQUENCE FROM N.A.
RA MEDLINE=97480713; PubMed=9339358;
RA Hagg P.M., Horelli-Kuitunen N., Eklund L., Palotie A.,
RA Pihlajaniemi T.;
RT "Cloning of mouse type XV collagen sequences and mapping of the
RT corresponding gene to 4B1-3. Comparison of mouse and human alpha 1
RT (XV) collagen sequences indicates divergence in the number of small
RT collagenous domains."
RT Genomics 45:31-41(1997).
RL EMBL; AF011450; AAC53387.1; -.
RL HSSP; P39061; 1KOE.
DR MGI; MGI:88449; COL15a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1367 AA; 140525 MW; A483A1254AF3AEBC CRC64;

Query Match
Best Local Similarity 57.7%; Score 515; DB 11; Length 1367;
Matches 98; Conservative 26; Mismatches 40; Indels 4; Gaps 1;

QY 1 VALNSPLSGMGIRGADFOCFQOARAVAGLAGTFRAFLVSLRDLVSVRRADRAAVPIV 60
DB 1200 VALNTVAAGDIR---ADFCFQOARAAAGLSTFRALVSLHLDLSTVAVKARFGPIV 1255

QY 61 NLKDELLFSEWELFFSGSEBPLKRGARIFSPDGKDVLRHPTWPKSVHSGSDPNRRLLTE 120
DB 1256 NLKGOVLFNNWDSIFSGDGQFNTHTPIYGFDRGVMTDPSWPKVYVHGSNPHGVRLVD 1315

QY 121 SYCETWRTSPATGQASLLGGRLLGQSAASCHHAYIVL 168
DB 1316 KYCAARITTMAYTGFPASPLSTKILDKAKYSCANRLIVLCIENSFMT 1363
```

```
RESULT 12
ID 09EOD9 PRELIMINARY; PRT; 1367 AA.
AC 09EOD9;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Type XV collagen.
GN COL15A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
[1]
RX SEQUENCE FROM N.A.
RA STRAIN=129/SV.
RX MEDLINE=97480713; PubMed=9339358;
RA Hagg P.M., Horelli-Kuitunen N., Eklund L., Palotie A.,
RA Pihlajaniemi T.;
RT "Cloning of mouse type XV collagen sequences and mapping of the
RT corresponding gene to 4B1-3. Comparison of mouse and human alpha 1
RT (XV) collagen sequences indicates divergence in the number of small
RT collagenous domains."
RT Genomics 45:31-41(1997).
RL Matrix Biol. 19:489-500(2000).
[2]
RX SEQUENCE FROM N.A.
RX STRAIN=129/SV;
RX MEDLINE=20522048; PubMed=11068203;
RA Eklund L., Wuona A., Lileard J., Pihlajaniemi T.;
RA "Structure of the mouse type XV collagen gene, COL15a1, comparison
RT with the human COL15A1 gene and functional analysis of the promoters
RT of both genes."
RT Matrix Biol. 19:489-500(2000).
DR EMBL; AF261131; AAG27545.1; JOINED.
DR EMBL; AF261109; AAG27545.1; JOINED.
DR EMBL; AF261110; AAG27545.1; JOINED.
DR EMBL; AF261111; AAG27545.1; JOINED.
DR EMBL; AF261112; AAG27545.1; JOINED.
DR EMBL; AF261113; AAG27545.1; JOINED.
DR EMBL; AF261114; AAG27545.1; JOINED.
DR EMBL; AF261115; AAG27545.1; JOINED.
DR EMBL; AF261116; AAG27545.1; JOINED.
DR EMBL; AF261117; AAG27545.1; JOINED.
DR EMBL; AF261118; AAG27545.1; JOINED.
DR EMBL; AF261119; AAG27545.1; JOINED.
DR EMBL; AF261120; AAG27545.1; JOINED.
DR EMBL; AF261121; AAG27545.1; JOINED.
DR EMBL; AF261122; AAG27545.1; JOINED.
DR EMBL; AF261123; AAG27545.1; JOINED.
DR EMBL; AF261124; AAG27545.1; JOINED.
DR EMBL; AF261125; AAG27545.1; JOINED.
DR EMBL; AF261126; AAG27545.1; JOINED.
DR EMBL; AF261127; AAG27545.1; JOINED.
DR EMBL; AF261128; AAG27545.1; JOINED.
DR EMBL; AF261129; AAG27545.1; JOINED.
DR EMBL; AF261130; AAG27545.1; JOINED.
DR HSSP; P39061; 1KOE.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1367 AA; 140454 MW; 40F259A3CD47C982 CRC64;

Query Match
Best Local Similarity 57.3%; Score 512; DB 11; Length 1367;
Matches 98; Conservative 25; Mismatches 41; Indels 4; Gaps 1;
```

QY 1 VALNSPLSGNRGRGADPFCFOQARAAGLAGTFRAPLSSRLQDLVSYVRRAAAYPIV 60
 DB 1200 VALNTPTVAAGDIR----ADFCFOQARAAGLAGTFRAPLSSRLQDLVSYVRRAAAYPIV 1255
 QY 61 NLKQVLFPMNNDISFSGGGGPNTHIRPIYSPDGRDVMTPDPSWPKVWHGSPHYGLVLD 120
 DB 1256 NLKQVLFPMNNDISFSGGGGPNTHIRPIYSPDGRDVMTPDPSWPKVWHGSPHYGLVLD 1315
 QY 121 SYCTWRTAPASATGQASLLGRLGQSAASHHAYIVLCIENSFMT 168
 DB 1316 KYCAMRTDPAVTGAFASPLSTGKILDKAKYSCANRLIVLCIENSFMT 1363
 RESULT 13
 QYV4M4 PRELIMINARY; PRT; 1368 AA.
 ID QYV4M4
 DT 01-NOV-1999 (T-REMBLrel. 12, Created)
 DT 01-NOV-1999 (T-REMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)
 DE Type XV collagen.
 GN COL15A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 ON NCBI_TaxId=9606;
 RN [1]
 RP MEDLINE=94148920; PubMed=8106446;
 RA Kivirikko S., Heinämäki P., Rehn M., Honkanen N., Myers J.C.,
 RA Phlajajaniemi T.;
 RT "Primary structure of the alpha 1 chain of human type XV collagen and
 RT exon-intron organization of the alpha 1 region of the corresponding gene.";
 RL J. Biol. Chem. 269:4773-4779(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98316357; PubMed=9653385;
 RA Hägg P.M., Muona A., Lietard J., Kivirikko S., Phlajajaniemi T.;
 RT "Complete exon-intron organization of the human gene for the alpha
 RT chain of type XV collagen (COL15A1) and comparison with the homologous
 RT COL18A1 gene.";
 RL J. Biol. Chem. 273:17824-17831(1998).
 DR EMBL; L25280; AAC78500.1; -
 DR EMBL; L25280; AAC78500.1; JOINED.
 DR EMBL; AF052956; AAC78500.1; JOINED.
 DR EMBL; AF052957; AAC78500.1; JOINED.
 DR EMBL; AF052958; AAC78500.1; JOINED.
 DR EMBL; AF052959; AAC78500.1; JOINED.
 DR EMBL; AF052960; AAC78500.1; JOINED.
 DR EMBL; AF052961; AAC78500.1; JOINED.
 DR EMBL; AF052962; AAC78500.1; JOINED.
 DR EMBL; AF052963; AAC78500.1; JOINED.
 DR EMBL; AF052964; AAC78500.1; JOINED.
 DR EMBL; AF052965; AAC78500.1; JOINED.
 DR EMBL; AF052966; AAC78500.1; JOINED.
 DR EMBL; AF052967; AAC78500.1; JOINED.
 DR EMBL; AF052968; AAC78500.1; JOINED.
 DR EMBL; AF052969; AAC78500.1; JOINED.
 DR EMBL; AF052970; AAC78500.1; JOINED.
 DR EMBL; AF052971; AAC78500.1; JOINED.
 DR EMBL; AF052972; AAC78500.1; JOINED.
 DR EMBL; AF052973; AAC78500.1; JOINED.
 DR EMBL; AF052974; AAC78500.1; JOINED.
 DR EMBL; AF052975; AAC78500.1; JOINED.
 DR EMBL; L25285; AAC78500.1; JOINED.
 DR EMBL; L25284; AAC78500.1; JOINED.
 DR EMBL; L25283; AAC78500.1; JOINED.
 DR EMBL; L25282; AAC78500.1; JOINED.
 DR EMBL; L25281; AAC78500.1; JOINED.
 DR HSSP; P33061.1KOB.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; Pf01391; Collagen; 5.

DR Pfam; PF02210; TSPN; 1.
 DR SMART; SM00282; Lsmg; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Collagen.
 SQ SEQUENCE 1368 AA; 141757 MW; 96828E45E847194B CRC64;
 Query Match 56.6%; Score 505; DB 4; Length 1368;
 Best Local Similarity 56.9%; Pref. No. 2e-42;
 Matches 95; Conservative 28; Mismatches 40; Indels 4; Gaps 1;
 QY 2 ALNSPLSGNRGRGADPFCFOQARAAGLAGTFRAPLSSRLQDLVSYVRRAAAYPIV 61
 DB 1222 ALNMPVSGDIR----ADFCFOQARAAGLAGTFRAPLSSRLQDLVSYVRRAAAYPIV 1277
 QY 62 LKQVLFPMNNDISFSGGGGPNTHIRPIYSPDGRDVMTPDPSWPKVWHGSPHYGLVLD 121
 DB 1278 LKQVLFPMNNDISFSGGGGPNTHIRPIYSPDGRDVMTPDPSWPKVWHGSPHYGLVLD 1337
 QY 122 YCETWRTAPASATGQASLLGRLGQSAASHHAYIVLCIENSFMT 168
 DB 1338 YCEAMRTDPAVTGAFASPLSTGKILDKAKYSCANRLIVLCIENSFMT 1364
 RESULT 14
 QYV509 PRELIMINARY; PRT; 581 AA.
 ID QYV509
 DT 01-MAY-2000 (T-REMBLrel. 13, Created)
 DT 01-MAY-2000 (T-REMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)
 DE CG8645 protein.
 GN CG8645.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN=BERKLEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brodeur P., Broxtie P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegang C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,
 RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wesserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AF003560; AAF50621.1; -.
DR HSSP; P39061; IKOR.
DR PiyBase; FBgn0035732; CG8645.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR002086; PPTA.
DR Pfam; PF01391; Collagen; 3.
DR ProDom; PD000007; Collagen; 1.
DR PROSITE; PS00904; PPTA; UNKNOWN 1.
DR PROSITE; PS00904; PPTA; UNKNOWN 1.
SQ SEQUENCE 581 AA; 60772 MW; 19EC1E48CB47FE7 CRC64;

Query Match 46.9%; Score 418.5; DB 5; Length 581;
Best Local Similarity 49.4%; Pred. No. 4.1e-34;
Matches 80; Conservative 28; Mismatches 53; Indels 1; Gaps 1;

QY 2 ALNSPLSGMGKIRGADPFCQARAVAGLAFRAFLSSRLQDLVSIYRRADRAAVPIV 61
DB 379 ALNBPSTGDLQIRGADPACRGGRAGLGTFRKAFSSRVQNDLTVRAPDR-DLEVN 437
QY 62 LKDELFPSEWALFSGSEGLPKGARIFFDQKDVLRHPTWPKSVHSGSDPNGRRLTES 121
DB 438 TRGDVLFNSWKGIFNGCGFFSQAPRIYFSFGKVMWTDSTPMKVMHSGSLPNGERSMDT 497
QY 122 YCETWTEAPSATGQASSILGRLGQSAASCHNAVIVLCIE 163
DB 498 YCDAMHSGDHLKGFASNLDPGHLKLEQKQSCSKLIILCIE 539

RESULT 15
017866 PRELIMINARY; PRT; 650 AA.
AC 017866
DT 01-JUN-1998 (TrEMBLrel. 05, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CLE-1C protein (C36B1.1c protein).
GN CLE-1 OR C36B1.1C.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21157401; PubMed=11257122;
RA Ackley B.D., Crew J.R., Blamaa H., Pihlajaniemi T., Kuo C.J.,
Kramer J.M.;
RT "The NCI/endostatin domain of Caenorhabditis elegans type XVIII
RT collagen affects cell migration and axon guidance";
RL J. Cell Biol. 152:1219-1232(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL; AF164959; AAD47825.1; -.
DR EMBL; 281079; CAB03084.1; -.
DR HSSP; P39061; IKOR.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; Collagen; 2.
SQ SEQUENCE 650 AA; 69596 MW; 6CF29BD9C16B170E CRC64;

Query Match 40.0%; Score 357; DB 5; Length 650;

Best Local Similarity 44.0%; Pred. No. 8.8e-28;
Matches 73; Conservative 24; Mismatches 63; Indels 6; Gaps 4;
QY 1 VALNSPLSGMGKIRGADPFCQARAVAGLAFRAFLSSRLQDLVSIYRRADRAAVPIV 60
DB 468 IALSQPSGNLHGLRGADLQCYREARAGYTTFRMLSSNVQDLVRIHVSVD-PDTTV 526
QY 61 NLKDELFPSEWALFSGSEGLPKGARIFFDQKDVLRHPTWPKSVHSGSDPNGRRLTE 120
DB 527 NVAGHHILFPMRSFVNVGAQ--NMPHAKLPSFDHVDLNDSRWPKKVMHSGSKGGIR-AE 563
QY 121 YCETWTEAPSATGQASSILGRLGQSAAS--CHNAVIVLCIE 164
DB 584 QYCDGWRADSLTSLAGHISNTSIFQSGSEKCNKLVLCVEN 629

Search completed: February 11, 2003, 20:42:38
Job time : 47 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:40:49 ; Search time 14 Seconds

(without alignments)
357,278 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893
Sequence: 1 VALNSPLSGMGRIGRADPQ.....ASCHHAYIVLCIENSFMTAS 170Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	893	100.0	178	4	US-09-315-689-5
2	893	100.0	182	4	US-09-561-500-14
3	893	100.0	182	4	US-09-561-108-14
4	893	100.0	182	4	US-09-315-689-3
5	893	100.0	182	4	US-09-561-526-14
6	893	100.0	183	4	US-09-206-059-2
7	778	87.1	191	4	US-09-561-500-13
8	778	87.1	191	4	US-09-561-108-13
9	778	87.1	191	4	US-09-561-526-13
10	775	86.8	195	1	US-08-159-784-2
11	734	82.2	185	3	US-08-985-526-36
12	491	55.0	191	1	US-08-159-784-3
13	160	17.9	35	3	US-09-046-985-2
14	160	17.9	32	4	US-09-474-743-2
15	101	11.3	22	3	US-09-046-985-7
16	101	11.3	22	4	US-09-474-743-7
17	94	10.5	16	4	US-09-385-442-32
18	76	8.5	578	2	US-08-653-740-3
19	76	8.5	578	2	US-09-073-594-3
20	76	8.5	578	3	US-09-275-925-3
21	76	8.5	636	1	US-08-553-740-5
22	76	8.5	636	2	US-09-073-594-5
23	76	8.5	636	3	US-09-275-925-5
24	74	8.3	256	1	US-07-906-349A-8
25	74	8.3	256	1	US-08-167-035-4
26	74	8.3	256	1	US-08-208-887A-4
27	74	8.3	256	2	US-08-539-005-4

28	74	8.3	256	4	US-09-280-598-8	Sequence 8, App11
29	73	8.2	311	3	US-08-987-743-2	Sequence 2, App11
30	73	8.2	435	3	US-08-733-360A-1	Sequence 1, App11
31	73	8.2	435	3	US-08-733-360A-3	Sequence 3, App11
32	73	8.2	435	3	US-08-987-743-6	Sequence 6, App11
33	73	8.2	435	3	US-08-987-743-15	Sequence 15, App11
34	73	8.2	435	4	US-08-916-935-1	Sequence 1, App11
35	72	8.1	435	4	US-08-916-935-3	Sequence 3, App11
36	72	8.1	304	4	US-08-630-915A-28	Sequence 28, App1
37	70	7.8	210	4	US-09-188-930-148	Sequence 148, App
38	70	7.8	953	4	US-09-245-281-43	Sequence 43, App1
39	70	7.8	953	4	US-09-207-352B-43	Sequence 43, App1
40	70	7.8	6095	4	US-09-144-085-2	Sequence 2, App11
41	69	7.7	190	3	US-09-046-985-15	Sequence 15, App1
42	69	7.7	190	4	US-09-474-743-15	Sequence 15, App1
43	69	7.7	313	4	US-09-347-803-6	Sequence 6, App11
44	69	7.7	327	6	5171684-6	Patent No. 5171684
45	68.5	7.7	1289	1	US-07-876-280-4	Sequence 4, App11

ALIGNMENTS

```
RESULT 1
US-09-315-689-5
; Sequence 5, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-5

Query Match      100.0%, Score 893; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 3,1e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGRADPQCFQQAARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAPVY 60
DB 9 VALNSPLSGMGRIGRADPQCFQQAARAVGLAGTFRAFLSSRLQDLVSIYRRADRAAPVY 68
QY 61 NLKDELLPFWSEALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWVGSDPNGRRLTR 120
DB 69 NLKDELLPFWSEALFSGSGPLKPGARIFSPGKDVLRHPTWPKSVWVGSDPNGRRLTR 128
QY 121 SYCETWRTPASATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
DB 129 SYCETWRTPASATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 178

RESULT 2
US-09-561-500-14
; Sequence 14, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip B. Thorpe
; APPLICANT: Rolf A. Brecken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
```

```
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 182
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-500-14
```

```
Query Match          100.0%; Score 893; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
    |||||
Db 13 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 120
    |||||
Db 73 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
    |||||
Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182
```

RESULT 3

```
US-09-561-108-14
/ Sequence 14, Application US/09561108
/ Patent No. 6342221
/ GENERAL INFORMATION:
/ APPLICANT: Philip E. Thorpe
/ APPLICANT: Rolf A. Brekken
/ TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
/ FILE REFERENCE: 4001.002584
/ CURRENT APPLICATION NUMBER: US/09/561.108
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/131,432
/ PRIOR FILING DATE: 1999-04-28
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 182
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-108-14
```

```
Query Match          100.0%; Score 893; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
    |||||
Db 13 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 120
    |||||
Db 73 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
    |||||
Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182
```

RESULT 4

```
US-09-315-689-3
/ Sequence 3, Application US/09315689
/ Patent No. 6346510
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Folkman, Judah
/ APPLICANT: O'Reilly, Michael
/ TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
/ FILE REFERENCE: 05213-0229
/ CURRENT APPLICATION NUMBER: US/09/315,689
/ CURRENT FILING DATE: 1999-05-20
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 3
/ LENGTH: 182
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-315-689-3
```

```
Query Match          100.0%; Score 893; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
    |||||
Db 13 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 120
    |||||
Db 73 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
    |||||
Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182
```

RESULT 5

```
US-09-561-526-14
/ Sequence 14, Application US/09561526
/ Patent No. 6416758
/ GENERAL INFORMATION:
/ APPLICANT: Philip E. Thorpe
/ APPLICANT: Rolf A. Brekken
/ TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
/ FILE REFERENCE: 4001.002586
/ CURRENT APPLICATION NUMBER: US/09/561,526
/ CURRENT FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/131,432
/ PRIOR FILING DATE: 1999-04-28
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 182
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-526-14
```

```
Query Match          100.0%; Score 893; DB 4; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
    |||||
Db 13 VALNSPLSGMRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 120
    |||||
Db 73 NLKDELLFPSEWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHSDDNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
    |||||
Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182
```

RESULT 6

US-09-206-059-2
Sequence 2, Application US/09206059
Patent No. 6201104
GENERAL INFORMATION:
APPLICANT: Macdonald, Nicholas
APPLICANT: Sam, Kim Lee
TITLE OF INVENTION: Angiogenesis-inhibiting Protein Binding Peptides and
FILE REFERENCE: 05213-0370
CURRENT APPLICATION NUMBER: US/09/206,059
CURRENT FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 183
TYPE: PRF
ORGANISM: Homo sapiens
US-09-206-059-2

Query Match 100.0%; Score 893; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 3, 2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 72
QY 61 NLKDELLFSPWEALFSGSEGPLKPGARIFSPDGKDYLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFSPWEALFSGSEGPLKPGARIFSPDGKDYLRHPTWPKSVHMGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHAYIVLCIENSPMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASHAYIVLCIENSPMTAS 182

RESULT 7
US-09-561-500-13
Sequence 13, Application US/09561500
Patent No. 6342219
GENERAL INFORMATION:
APPLICANT: Phillip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002500
CURRENT APPLICATION NUMBER: US/09/561,500
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 191
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-500-13

Query Match 87.1%; Score 778; DB 4; Length 191;
Best Local Similarity 85.8%; Pred. No. 3, 8e-87;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 20 VALNTPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 79
QY 61 NLKDELLFSPWEALFSGSEGPLKPGARIFSPDGKDYLRHPTWPKSVHMGSDPNGRRLTE 120
DB 80 NLKDEVLSPSWDSLFGSGQGLQPGARIFSPDRDYLRRHAPWPKSVHMGSDPSGRRLME 139
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHAYIVLCIENSPMTA 169

DB 140 SYCETWRTTGTATGQASSLLGRLLEQKASCHNSIYIVLCIENSPMTS 188

RESULT 8
US-09-561-108-13
Sequence 13, Application US/09561108
Patent No. 6342221
GENERAL INFORMATION:
APPLICANT: Phillip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002584
CURRENT APPLICATION NUMBER: US/09/561,108
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 191
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-108-13

Query Match 87.1%; Score 778; DB 4; Length 191;
Best Local Similarity 85.8%; Pred. No. 3, 8e-87;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 20 VALNTPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 79
QY 61 NLKDELLFSPWEALFSGSEGPLKPGARIFSPDGKDYLRHPTWPKSVHMGSDPNGRRLTE 120
DB 80 NLKDEVLSPSWDSLFGSGQGLQPGARIFSPDRDYLRRHAPWPKSVHMGSDPSGRRLME 139
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHAYIVLCIENSPMTA 169
DB 140 SYCETWRTTGTATGQASSLLGRLLEQKASCHNSIYIVLCIENSPMTS 188

RESULT 9
US-09-561-526-13
Sequence 13, Application US/09561526
Patent No. 6416758
GENERAL INFORMATION:
APPLICANT: Phillip E. Thorpe
APPLICANT: Rolf A. Brekken
TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
FILE REFERENCE: 4001.002586
CURRENT APPLICATION NUMBER: US/09/561,526
CURRENT FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/131,432
PRIOR FILING DATE: 1999-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 191
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-526-13

Query Match 87.1%; Score 778; DB 4; Length 191;
Best Local Similarity 85.8%; Pred. No. 3, 8e-87;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 20 VALNTPLSGMGRIGADPFCFOQARAVGLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 79

[illegible]

RESULT 10
US-08-159-784-2

Query Match	86.8%;	Score 775;	DB 1;	Length 195;
Best Local Similarity	85.2%;	Pred. No. 9.2e-87;		
Matches 144; Conservative	14;	Mismatches	11;	Gaps 0

[illegible]

RESULT 11
US-08-985-526-36
; Sequence 36; Application US/08985526

Patent No. 6080728
GENERAL INFORMATION:
APPLICANT: Mixson, James A
TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENET
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Connolly, Bove, Lodge, & Hutz
STREET: 1220 Market Street, P.O. Box 2207
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorow Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-985,526-36

[illegible]

RESULT 12
US-08-159-784-3
Sequence 3, Application US/08159784
Patent No. 5643783
GENERAL INFORMATION:
APPLICANT: Bjorn R. Olsen
TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Ps/2 Model 502 or 555X


```

US-09-046-985-7
: Sequence 7, Application US/090465985
: Patent No. 6121236
: GENERAL INFORMATION:
: APPLICANT: Ben-Sasson, Shmuel A.
: TITLE OF INVENTION: NOVEL MULTIVALENT LIGANDS WHICH MODULATE
: TITLE OF INVENTION: ANGIOGENESIS
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/046,985
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: CMCC-614
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781) 861-6240
: TELEFAX: (781) 861-9540
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 22 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 1
: OTHER INFORMATION: /label= modified aa
: OTHER INFORMATION: /note= "N-Acetyl-L-Threonine"
US-09-046-985-7

Query Match 11.3%; Score 101; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 TTRAFSSRLQDLYSTIVRRAD 53
DB 1 TTRAFSSRLQDLYSTIVRRAD 21

```

Search completed: February 11, 2003, 20:43:26
Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 20:42:44 ; Search time 30 Seconds

(Without alignments)
144.777 Million cell updates/sec

Title: US-09-171-607-1

Perfect score: 893
Sequence: 1 VALNSPLSGMGIRGADPQ.....ASCHHAYIVLCIENSFMTAS 170

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548976 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	893	100.0	178	US-10-131-241-60	Sequence 60, Appl
2	893	100.0	179	US-10-131-241-57	Sequence 57, Appl
3	893	100.0	182	US-10-131-241-54	Sequence 54, Appl
4	893	100.0	182	US-09-998-831-14	Sequence 14, Appl
5	893	100.0	183	US-10-080-797-1	Sequence 1, Appl
6	893	100.0	183	US-10-131-241-52	Sequence 52, Appl
7	893	100.0	183	US-09-873-676-2	Sequence 2, Appl
8	889	99.6	181	US-10-131-241-55	Sequence 55, Appl
9	885	99.1	180	US-10-131-241-56	Sequence 56, Appl
10	858	96.1	180	US-10-131-241-47	Sequence 47, Appl
11	783	87.7	184	US-10-131-241-49	Sequence 49, Appl
12	778	87.1	191	US-09-998-831-13	Sequence 13, Appl
13	778	87.1	207	US-10-080-797-3	Sequence 3, Appl
14	775	86.8	184	US-10-131-241-46	Sequence 46, Appl
15	734	82.2	185	US-10-036-869-36	Sequence 36, Appl
16	346	38.7	63	US-09-823-540A-1	Sequence 1, Appl
17	182	20.4	31	US-09-822-540A-2	Sequence 2, Appl
18	94	10.5	16	US-09-766-412-32	Sequence 32, Appl
19	85	9.5	1244	US-09-815-915-8	Sequence 8, Appl

20	77.5	8.7	418	10	US-09-927-602-5	Sequence 5, Appl
21	75	8.4	332	10	US-09-815-242-5659	Sequence 5659, Ap
22	75	8.4	346	10	US-09-815-242-12271	Sequence 12271, A
23	73.5	8.2	1653	10	US-09-741-669-402	Sequence 402, App
24	72	8.1	304	10	US-09-879-957-28	Sequence 28, Appl
25	72	8.1	636	9	US-10-000-776-12	Sequence 12, Appl
26	72	8.1	636	9	US-09-791-457-14	Sequence 14, Appl
27	70	7.8	953	9	US-10-118-984-43	Sequence 43, Appl
28	70	7.8	953	10	US-09-728-721-43	Sequence 43, Appl
29	68.5	7.7	1016	9	US-09-738-626-4363	Sequence 4363, Ap
30	68.5	7.7	1289	10	US-09-738-163-4	Sequence 4, Appl
31	68	7.6	200	9	US-10-118-984-11	Sequence 11, Appl
32	68	7.6	200	10	US-09-728-721-11	Sequence 11, Appl
33	68	7.6	200	12	US-10-105-931-11	Sequence 11, Appl
34	68	7.6	320	9	US-10-002-974-30	Sequence 30, Appl
35	68	7.6	320	12	US-10-014-269-30	Sequence 30, Appl
36	68	7.6	490	9	US-10-118-984-26	Sequence 26, Appl
37	68	7.6	490	10	US-09-728-721-26	Sequence 26, Appl
38	68	7.6	490	12	US-10-105-931-26	Sequence 26, Appl
39	68	7.6	953	9	US-10-118-984-8	Sequence 8, Appl
40	68	7.6	953	10	US-09-728-721-8	Sequence 8, Appl
41	68	7.6	953	12	US-10-105-931-8	Sequence 8, Appl
42	68	7.6	999	9	US-10-174-590-434	Sequence 434, App
43	68	7.6	999	9	US-10-176-758-434	Sequence 434, App
44	68	7.6	999	9	US-10-175-737-434	Sequence 434, App
45	68	7.6	999	9	US-10-173-706-434	Sequence 434, App

ALIGNMENTS

RESULT 1
US-10-131-241-60
; Sequence 60, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT FILING DATE: 2002-07-22
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-60

Query Match 100.0%; Score 893; DB 9; Length 178;
Best Local Similarity 100.0%; Pred. No. 4.7e-90;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADPQCFQARAVALGTFAPLSSRLQDIYSYRRADRAVPV 60
DB 9 VALNSPLSGMGIRGADPQCFQARAVALGTFAPLSSRLQDIYSYRRADRAVPV 68
QY 61 NIKDELLPFSWALPFGSGSGPLKPGARIFSPDGKDVLRHPTVPQKSVHMGSDPNGRRLTE 120
DB 69 NIKDELLPFSWALPFGSGSGPLKPGARIFSPDGKDVLRHPTVPQKSVHMGSDPNGRRLTE 128
QY 121 SYCETWRTAPASATGQASLLGGRLLGSGAASCHHAYIVLCIENSFMTAS 170
DB 129 SYCETWRTAPASATGQASLLGGRLLGSGAASCHHAYIVLCIENSFMTAS 178

RESULT 2
US-10-131-241-57
; Sequence 57, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-57

Query Match 100.0%; Score 893; DB 9; Length 179;
Best Local Similarity 100.0%; Pred. No. 4.7e-90;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 9 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 68
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 120
DB 69 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 128
QY 121 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 170
DB 129 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 178

RESULT 3
US-10-131-241-54
; Sequence 54, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-54

Query Match 100.0%; Score 893; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.8e-90;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 132
QY 121 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 170
DB 133 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 182

RESULT 4
US-09-998-831-14
; Sequence 14, Application US/09998831
; Patent No. US2002019153A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-998-831-14

Query Match 100.0%; Score 893; DB 10; Length 182;
Best Local Similarity 100.0%; Pred. No. 4.8e-90;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVHGSDPNGRRLTE 132
QY 121 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 170
DB 133 SYCETWTEPASPATGQASSLLGRLGQSAASHAYIVLCIENSFMTAS 182

RESULT 5
US-10-080-797-1
; Sequence 1, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Camacho-Ibar, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 183

TYPE: prt
ORGANISM: Human
US-10-080-797-1

Query Match
Best Local Similarity 100.0%; Score 893; DB 9; Length 183;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 72
QY 61 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 6

US-10-131-241-52
Sequence 52, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT FILING DATE: 2002-07-22
PRIOR FILING DATE: 1999-10-06
PRIOR FILING DATE: 1999-10-06
PRIOR FILING DATE: 1999-05-21
PRIOR FILING DATE: 1999-05-21
PRIOR FILING DATE: 1998-05-22
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 183
TYPE: prt
ORGANISM: Homo sapiens
US-10-131-241-52

Query Match
Best Local Similarity 100.0%; Score 893; DB 9; Length 183;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 72
QY 61 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 7

US-09-873-676-2
Sequence 2, Application US/09873676
Patent No. US20020077289A1
GENERAL INFORMATION:
APPLICANT: Macdonald, Nicholas J.
APPLICANT: Sim, Kim L.
TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
FILE REFERENCE: 05213-0378 (43170-259333)

CURRENT APPLICATION NUMBER: US/09/873.676
CURRENT FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: US 60/209,065
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: US 60/289,387
PRIOR FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 183
TYPE: prt
ORGANISM: Homo sapiens
US-09-873-676-2

Query Match
Best Local Similarity 100.0%; Score 893; DB 10; Length 183;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 72
QY 61 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 8

US-10-131-241-55
Sequence 55, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
FILE REFERENCE: 05213-0344 43170-271565
CURRENT FILING DATE: 2002-07-22
PRIOR FILING DATE: 1999-10-06
PRIOR FILING DATE: 1999-10-06
PRIOR FILING DATE: 1999-05-21
PRIOR FILING DATE: 1999-05-21
PRIOR FILING DATE: 1998-05-22
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55
LENGTH: 181
TYPE: prt
ORGANISM: Homo sapiens
US-10-131-241-55

Query Match
Best Local Similarity 99.6%; Score 889; DB 9; Length 181;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 60
DB 13 VALNSPLSGMGIRGADPQCFOQARAAGVLAGTFRAFLSSRLQDLYSIYRRADRAAVPIV 72
QY 61 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 120
DB 73 NLKDELLFSPSWALFSGSGEPLKPGARIFSPDGKDVLRHPTWPKSVHGSDPNGRRLTE 132
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTA 169
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTA 181

RESULT 9
US-10-131-241-56
Sequence 56, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
TITLE OF INVENTION: Compositions and Methods for Inhibiting Angiogenesis Using Cancer Markers
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 56
LENGTH: 180
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-241-56

Query Match 99.1%; Score 885; DB 9; Length 180;
Best Local Similarity 100.0%; Pred. No. 3,5e-89;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 132
QY 121 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 168
DB 133 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 180

RESULT 10
US-10-131-241-47
Sequence 47, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
TITLE OF INVENTION: Compositions and Methods for Inhibiting Angiogenesis Using Cancer Markers
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 47
LENGTH: 180
TYPE: PRT
ORGANISM: Rhesus monkey
US-10-131-241-47

Query Match 96.1%; Score 858; DB 9; Length 180;
Best Local Similarity 95.8%; Pred. No. 3,1e-86;
Matches 161; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 132
QY 121 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 168
DB 133 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 180

RESULT 11
US-10-131-241-49
Sequence 49, Application US/10131241
Publication No. US20030012792A1
GENERAL INFORMATION:
APPLICANT: Holaday, John W.
APPLICANT: Fortier, Anne H.
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
TITLE OF INVENTION: Compositions and Methods for Inhibiting Angiogenesis Using Cancer Markers
FILE REFERENCE: 05213-0344 43170-271565
CURRENT APPLICATION NUMBER: US/10/131,241
PRIOR FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/413,049
PRIOR FILING DATE: 1999-10-06
PRIOR APPLICATION NUMBER: US 09/316,802
PRIOR FILING DATE: 1999-05-21
PRIOR APPLICATION NUMBER: US 60/086,586
PRIOR FILING DATE: 1998-05-22
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 49
LENGTH: 184
TYPE: PRT
ORGANISM: Canine sp.
US-10-131-241-49

Query Match 87.7%; Score 783; DB 9; Length 184;
Best Local Similarity 85.8%; Pred. No. 4,8e-78;
Matches 145; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFCFOQARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVHMGSDPNGRRLTE 132
QY 121 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 168
DB 133 SYCETWRTPEASATGQASSLLGRLGQSAASCHHAYIVLCIENSFMT 180

RESULT 12
US-09-998-831-13
Sequence 13, Application US/09998831
Patent No. US20020119153A1
GENERAL INFORMATION:
APPLICANT: Philip E. Thorpe
APPLICANT: Rolf A. Brecken
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
TITLE OF INVENTION: INHIBITING VEGF
FILE REFERENCE: 4001,002584
CURRENT APPLICATION NUMBER: US/09/998,831
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 09/561,108
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-998-831-13

```

```

Query Match      87.1%; Score 778; DB 10; Length 191;
Best Local Similarity 85.8%; Pred. No. 1.8e-77;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1 VALNSPLSGMGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLYSIVRRADRAVPY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 20 VALNTPLSGMGIRGADFOCFQOARAVGLSTFRFLSSRLQDLYSIVRRADRGSPY 79
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NLKDELLPSPWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 80 NLKDEVLPSPWDSLFGSGQGLQPGARIFSPDGRDVLRRHAWPKSVWHGSDPSGRRLME 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SYCETWRTTAPSATGQASSLLGRLGQSAASHAYIVLCIENSFMTA 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 140 SYCETWRTTGTGATGQASSLLSGRLLEQKAASCHNSYIVLCIENSFMTS 188
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 13
US-10-080-797-3
; Sequence 3, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; TITLE OF INVENTION: NEOVASCULARIZATION
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Mouse
US-10-080-797-3

```

```

Query Match      87.1%; Score 778; DB 9; Length 207;
Best Local Similarity 85.8%; Pred. No. 2e-77;
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1 VALNSPLSGMGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLYSIVRRADRAVPY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 36 VALNTPLSGMGIRGADFOCFQOARAVGLSTFRFLSSRLQDLYSIVRRADRGSPY 95
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NLKDELLPSPWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 96 NLKDEVLPSPWDSLFGSGQGLQPGARIFSPDGRDVLRRHAWPKSVWHGSDPSGRRLME 155
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SYCETWRTTAPSATGQASSLLGRLGQSAASHAYIVLCIENSFMTA 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 156 SYCETWRTTGTGATGQASSLLSGRLLEQKAASCHNSYIVLCIENSFMTS 204
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 14
US-10-131-241-46
; Sequence 46, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565

```

```

; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-46

```

```

Query Match      86.8%; Score 775; DB 9; Length 184;
Best Local Similarity 85.2%; Pred. No. 3.6e-77;
Matches 144; Conservative 14; Mismatches 11; Indels -0; Gaps 0;

```

```

QY 1 VALNSPLSGMGIRGADFOCFQOARAVGLAGTFRAFLSSRLQDLYSIVRRADRAVPY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 13 VALNTPLSGMGIRGADFOCFQOARAVGLSTFRFLSSRLQDLYSIVRRADRGSPY 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NLKDELLPSPWEALFSGSEGPLKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 73 NLKDEVLPSPWDSLFGSGQGLQPGARIFSPDGRDVLRRHAWPKSVWHGSDPSGRRLME 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SYCETWRTTAPSATGQASSLLGRLGQSAASHAYIVLCIENSFMTA 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 133 SYCETWRTTGTGATGQASSLLSGRLLEQKAASCHNSYIVLCIENSFMTS 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 15
US-10-036-869-36
; Sequence 36, Application US/10036869
; Patent No. US2002015156A1
; GENERAL INFORMATION:
; APPLICANT: Mlxson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESS: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/036,869
; FILING DATE: 29-Nov. US2002015156A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid

```

```

;          TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-10-036-869-36

Query Match      82.2%; Score 734; DB 12; Length 185;
Best Local Similarity 82.4%; Pred. No. 1,1e-72;
Matches 140; Conservative 14; Mismatches 14; Indels 2; Gaps 2;

QY 1 VALNSPLSGMGRGIRGADFOCFQOARAVAGLAGTFRAFLSRLODLYSIVRRADRAVPIY 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 VALNTPLSGMGRGIRGADFOCFENNAR-VGLSGTFRAFLSRLODLYSIVRRADRGVPIY 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 -NLKDELLFPSEWALFSGSEGPLKPGARIFSGDKVDVLRHPTWPQKSVWGHGSDPNGRRLT 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 QNLRDEVLSFMSDSLFSGSGQLOPGARIFSFDRDVLRRHPAMPORSVWGHGSDPSGRRLM 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 120 ESYCETWRTETPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFPMTA 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 ESYCETWRTETTGATGQASSLLSGRLLEQRAASCHDSYIVLCIENSFPMTS 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: February 11, 2003, 20:47:34
Job time : 31 secs